

THE SEDONA CONFERENCE JOURNAL®

V o l u m e 1 5 ❖ F a l l 2 0 1 4

A R T I C L E S

In Memoriam: Richard G. Braman Craig W. Weinlein

Responding to the Government's Civil Investigations. David C. Shonka

What's the Problem with Google? Daniel R. Shulman

Technology-Assisted Review: The Judicial Pioneers
. Paul E. Burns & Mindy M. Morton

**The Sedona Conference Commentary on Patent Damages
& Remedies**. The Sedona Conference

The Sedona Conference Commentary on Information Governance
. The Sedona Conference

**The Sedona Conference Database Principles Addressing the
Preservation & Production of Databases & Database
Information in Civil Litigation** The Sedona Conference

**The Sedona Conference Best Practices Commentary on the
Use of Search & Information Retrieval Methods in E-Discovery**
. The Sedona Conference

**The Sedona Conference Commentary
on Achieving Quality in the E-Discovery Process** . . . The Sedona Conference

**The Sedona Conference Glossary: E-Discovery & Digital
Information Management (Fourth Edition)** The Sedona Conference



ANTITRUST LAW, COMPLEX LITIGATION
AND INTELLECTUAL PROPERTY RIGHTS

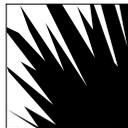
THE SEDONA CONFERENCE JOURNAL®

VOLUME 15



FALL 2014

The
sedona
conference®



The Sedona Conference Journal® (ISSN 1530-4981) is published on an annual basis, containing selections from the preceding year's Conferences and Working Group efforts.

The Journal is available on a complementary basis to courthouses and public law libraries and by subscription to others (\$95; \$45 for Conference participants and Working Group members).

Send us an email (info@sedonaconference.org) or call (1-602-258-4910) to order or for further information. Check our website for further information about our Conferences, Working Groups, and publications: www.thesedonaconference.org.

Comments (strongly encouraged) and requests to reproduce all or portions of this issue should be directed to:

The Sedona Conference, 5150 North 16th Street, Suite A-215,
Phoenix, AZ 85016 or call 1-602-258-4910; fax 602-258-2499; email info@sedonaconference.org.

The Sedona Conference Journal® designed by MargoBDesign.com – mbraman@sedona.net

Cite items in this volume to “15 Sedona Conf. J. _____ (2014).”

Copyright 2014, The Sedona Conference.
All Rights Reserved.

PUBLISHER'S NOTE

Welcome to Volume 15 of *The Sedona Conference Journal* (ISSN 1530-4981), an annual collection of articles originally presented at our Conferences, and commentaries prepared by our Working Groups, over the past year. The Sedona Conference was founded in 1997 to provide a forum for advanced dialogue by the nation's leading attorneys, academics and jurists of cutting-edge issues of law and policy in the areas of antitrust, intellectual property rights, and complex litigation. We host Regular Season Conferences, international programmes, TSCI conferences and several Working Group meetings each year, providing unique and rewarding opportunities to seriously explore the boundaries of various areas of the law with those who are creating them. This volume of the *Journal* contains two articles from The Sedona Conference Institute (TSCI) Program (March 2014), one article from our regular season conference on antitrust law and litigation (July 2014) and six Working Group commentaries that have been published since the printing of *The Sedona Conference Journal*, Volume 14, in 2013.

We hope that you will find that the papers in this *Journal* reflect the same mix of theory and experience found at our Conferences and Working Group Meetings, including the creativity and constructive irreverence required to challenge traditional thinking. The views expressed herein are those of the authors, and we encourage the submission of counterpoint pieces. Submissions can be sent electronically to info@sedonaconference.org, or by mail to The Sedona Conference, 5150 North 16th Street, Suite A-215, Phoenix, AZ 85016, USA. If you are interested in participating in one of our Regular Season Conferences, our TSCI conferences or international programmes, or in joining our Working Group Series, please visit our website for further information (www.thesedonaconference.org).

Craig Weinlein
Executive Director
The Sedona Conference
November 2014

The Sedona Conference gratefully acknowledges the substantial contributions of its Conference faculties, Working Group Series Sustaining and Annual Sponsors, participants, members and observers, and our Advisory Board members, whose volunteer efforts and contributions make The Sedona Conference a “thought-provoking and inspiring” experience providing content of immediate benefit to the Bench and Bar.

IN MEMORIAM

Richard G. Braman

Richard Braman was the visionary founder and Executive Director Emeritus of The Sedona Conference. His passing on June 9, 2014 was a loss to The Sedona Conference and to the legal profession.

Richard was a pathfinder and an agent of change. When he saw a need he found a solution. He found ways to foster cooperation and bridge gaps, and his natural leadership attracted others to join him in crafting solutions and finding a better way.

Richard was active as a leader and agent of change as early as high school, where he graduated at the top of his class. He was president of the thespians, the debate team, vice president of the student council, and president of the student federation. Richard created the federation as an alternative to a student council that was not responsive to the needs of students. He saw a need and, with the help of others, crafted a solution.

Richard attended the University of Minnesota, completed his course work in three years, and graduated *summa cum laude* in 1975. He then attended the University of Minnesota Law School and graduated *magna cum laude*. His professional career began in the antitrust group of the law firm of Brobeck, Phleger & Harrison in San Francisco. A year later, he realized that a large California law firm was not a good fit for his independence. He moved back to Minnesota and joined the boutique plaintiffs' class action firm of Opperman & Paquin.

In 1984, Richard took the first of two sabbaticals and opened Gabriel's, a Minnesota jazz club that quickly acquired national renown. He then resumed his law career and joined the law firm of Gray, Plant, Mooty, Mooty & Bennett (currently known as Gray Plant Mooty) in its plaintiffs' antitrust practice. He would co-chair the Antitrust Practice Group before taking his second sabbatical to create The Sedona Conference.

Richard founded The Sedona Conference in December 1997 because he was dissatisfied with traditional continuing legal education (CLE) programs. He knew there was a better way than "talking heads" delivering lectures to passive unengaged attendees who were collecting CLE credits while solving crossword puzzles or reading the paper. Richard wanted to replicate the college seminar, with a small group of people interested in the advanced subject matter and engaging in real dialogue on an issue. The result was The Sedona Conference regular season conferences on antitrust law, complex litigation, and intellectual property rights. From the beginning, these regular season conferences have been advanced level, dialogue-based mini sabbaticals with limited attendance for the nation's leading jurists, lawyers, and experts to examine cutting edge issues of law and policy. To encourage the dialogue leaders at the regular season conferences to create superior written materials to accompany and support the sessions at the conferences, Richard created *The Sedona Conference Journal* to publish the best papers from The Sedona Conference regular season conferences.

At the regular season conferences, leading practitioners were discussing possible solutions to cutting edge issues and wanted to follow-up by drafting principles, guidelines, or best practices to move the law forward. This naturally led to the formation of The Sedona Conference Working Groups. Working Group 1 produced *The Sedona Principles: Best Practices Recommendations & Principles Addressing Electronic Document Production* (March 2003 version), which had an

immediate and substantial impact. Within a few weeks they were cited by the Discovery Subcommittee of the Civil Rules Advisory Committee of the Judicial Conference of the United States. Less than a month later they were cited by Judge Shira Scheindlin of the United States District Court for the Southern District of New York in *Zubulake v. UBS Warburg*, the groundbreaking case on eDiscovery.

Because of Richard's energy and passion for the mission of The Sedona Conference, the organization has grown greatly in the past fifteen years. The Sedona Conference now has eleven Working Groups addressing issues in eDiscovery, antitrust, complex litigation, intellectual property rights, cyber security, and international electronic information management. Collectively, the Working Groups have published a total of 47 nonpartisan consensus commentaries, principles, guidelines, or best practices designed to move the law forward in a reasoned and just way. The newest Working Group is identifying and commenting on trends in data security and privacy law, in an effort to help organizations prepare for and respond to data breaches, and to assist attorneys and judicial officers in resolving questions of legal liability and damages.

As a result of Richard's work in the field of eDiscovery, in 2013 he was selected as one of the top six eDiscovery pioneers by *Law Technology News* and *The American Lawyer*. He was also named one of the 50 most innovative people in Big Law in the past 50 years by *The American Lawyer*. Richard also is remembered for his revolutionary work in fostering cooperation between opposing parties. Richard understood that cooperation, not obstruction, is the hallmark of a truly great and effective lawyer. *The Sedona Conference Cooperation Proclamation*, published in 2008, receives more judicial endorsements every year.

In my opinion, The Sedona Conference is the natural, if not inevitable, outgrowth of Richard being a pathfinder and an agent of change. Richard's personal principles and beliefs became the guiding principles for The Sedona Conference: professionalism, civility, an open mind, respect for the beliefs of others, thoughtfulness, reflection, and belief in a process based on civilized dialogue, not debate. The goodness and purity of the mission of The Sedona Conference combined with Richard's personal character, graciousness and friendly personality, attracted many of us, including me, to join Richard in supporting The Sedona Conference as an agent of change in the legal profession.

We best honor Richard and preserve his legacy by collectively carrying out the mission of The Sedona Conference to move the law forward in a reasoned and just way. In support of its mission, The Sedona Conference will continue to address unsettled and uncharted areas of the law, where the authoritative and influential guidance of The Sedona Conference can move the law forward. That is exactly what Richard would want.

Richard's portrait hangs in the hallway of The Sedona Conference office, a few steps from my office door. Every day I see Richard smiling as I walk by. I want to believe he is pleased with the work we are doing at The Sedona Conference, and proud of the organization he started. Thank you, Richard, for creating The Sedona Conference to act as an agent of change in the legal profession.

Craig W. Weinlein
Executive Director
The Sedona Conference
Phoenix, AZ

Joseph M. Alioto, Esq.
Alioto Law Firm, San Francisco, CA

Kevin F. Brady, Esq.
Redgrave LLP, Washington, DC

Elizabeth J. Cabraser, Esq.
Lieff Cabraser Heimann & Bernstein, San Francisco, CA

Professor Stephen Calkins, Esq.
Wayne State University Law School, Detroit, MI

The Hon. Justice Colin Campbell
Neeson Arbitration Chambers, Toronto, ON, Canada

The Hon. John L. Carroll (ret.)
Professor, Cumberland School of Law, Samford University, Birmingham, AL

Joe Cecil, Ph.D., J.D.
Federal Judicial Center, Division of Research, Washington, DC

Michael V. Ciresi, Esq.
Robins Kaplan Miller & Ciresi, Minneapolis, MN

The Hon. John Facciola
District of Columbia, Washington, DC

Prof. Steven S. Gensler
University of Oklahoma College of Law, Norman, OK

Michael D. Hausfeld, Esq.
Hausfeld LLP, Washington, DC

Prof. George A. Hay
Edward Cornell Professor of Law, Cornell Law School, Ithaca, NY

Hon. Katharine Sweeney Hayden
District of New Jersey, Newark, NJ

Ronald J. Hedges, Esq.
Ronald J. Hedges LLC, Hackensack, NJ

The Hon. Susan Illston
Northern District of California, San Francisco, CA

Allan Kanner, Esq.
Kanner & Whiteley, New Orleans, LA

Hon. Justice Gilles Letourneau
Federal Court of Appeal, Ottawa, Ontario

The Hon. J. Thomas Marten
U. S. District Court, District of Kansas, Wichita, KS

Hon. Paul R. Michel (ret.)
Alexandria, VA

Dianne M. Nast, Esq.
Roda & Nast P.C., Lancaster, PA

The Hon. Nan Nolan (ret.)
Chicago, IL

The Hon. Kathleen O'Malley
Federal Circuit Court of Appeals, Washington, DC

Vance K. Opperman, Esq.
Key Investment, Inc., Minneapolis, MN

The Hon. Andrew Peck
Southern District of New York, New York, NY

M. Laurence Popofsky, Esq.
Orrick LLP, San Francisco, CA

Jonathan M. Redgrave, Esq.
Redgrave LLP, Washington, DC

The Hon. James M. Rosenbaum (ret.)
Minneapolis, MN

Hon. Barbara Jacobs Rothstein
District Court for the District of Columbia, Washington, DC

Prof. Stephen A. Saltzburg
George Washington University Law School, Washington, D.C.

The Hon. Shira A. Scheindlin
Southern District of New York, New York, NY

The Hon. Craig Shaffer
District of Colorado, Denver, CO

Daniel R. Shulman, Esq.
Gray Plant Mooty, Minneapolis, MN

Robert G. Sterne, Esq.
Sterne Kessler Goldstein & Fox PLLC, Washington, DC

Dennis R. Suplee, Esq.
Schnader LLP, Philadelphia, PA

Prof. Jay Tidmarsh
Notre Dame Law School, Notre Dame, IN

Barbara E. Tretheway, Esq.
HealthPartners, Bloomington, MN

Hon. Ira B. Warshawsky
Meyer, Suozzi, English & Klein, P.C., Garden City, NY

The Hon. Carl J. West (ret.)
Los Angeles, CA

TABLE OF CONTENTS

Publisher's Note	i
In Memoriam: Richard G. Braman	ii
The Sedona Conference Advisory Board	iv
Responding to the Government's Civil Investigations	
David C. Shonka	1
What's the Problem with Google?	
Daniel R. Shulman	17
Technology-Assisted Review: The Judicial Pioneers	
Paul E. Burns & Mindy M. Morton	35
The Sedona Conference Commentary on Patent Damages & Remedies	
The Sedona Conference	53
The Sedona Conference Commentary on Information Governance	
The Sedona Conference	125
The Sedona Conference Database Principles Addressing the Preservation & Production of Databases & Database Information on Civil Litigation	
The Sedona Conference	171
The Sedona Conference Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery	
The Sedona Conference	217
The Sedona Conference Commentary on Achieving Quality in the E-Discovery Process	
The Sedona Conference	265
The Sedona Conference Glossary: E-Discovery & Digital Information Management (Fourth Edition)	
The Sedona Conference	305

THIS PAGE LEFT INTENTIONALLY BLANK.

RESPONDING TO THE GOVERNMENT'S CIVIL INVESTIGATIONS

*David C. Shonka*¹
Federal Trade Commission
Washington, D.C.

INTRODUCTION

Litigators often say that litigation is more about storytelling than it is about the facts. The goal is to put together a coherent, plausible, and sympathetic story that will grab the attention and interest of the audience and compel a conclusion favorable to the litigator's client. In contrast, government investigations are not about storytelling. They are about facts and the opinions and conclusions that can be drawn from the facts. More specifically, they are about obtaining a complete set of facts. Incomplete facts can lead to incorrect decisions. Incomplete facts can just as easily lead to law enforcement actions being filed that should not be filed as they can lead to important cases that should be filed not being filed. Either of these outcomes can be very expensive for the government, the parties, taxpayers, and the general public interest. This paper is about civil law enforcement investigations, the way the government conducts them, and the ways in which the parties to an investigation might appropriately deal with them.

A. Basics of Government Investigations

Government investigations are not like civil litigation. With respect to gathering facts, they differ from litigation in three major respects. First, only one party in an investigation gets to ask for documents and question witnesses. The government investigates so that it alone might collect information and decide whether to close an inquiry (as it does in most cases) or pursue further action, usually in the form of an administrative or judicial proceeding.

Second, investigations usually are conducted before any cause of action is identified, any claim or defense is asserted, or any court complaint is filed and served. This means the Federal Rules of Civil Procedure do not apply and discovery is not limited by the same notions of relevance that apply in litigation. While litigating parties may seek information relevant to a claim or defense of a party, the government may seek any information that is reasonably related to the scope of its investigation.²

1 The views expressed herein are solely those of the author and do not represent the views of the Federal Trade Commission, any individual Commissioner, or any other Commission employee. Preliminary versions of this article were presented at the 6th and 9th Annual Georgetown Advanced eDiscovery Institute programs held on December 12, 2009 and December 5, 2012 respectively, and at the Practising Law Institute: Government Investigations 2014 program held on March 26, 2014. The author appreciates the opportunity for dialogue that these venues permitted. The author also thanks Jeane A. Thomas, Crowell & Moring, LLP, for her very substantial suggestions and insights, especially with respect to the perspective of the private practitioner, and Jonathan Hill, FTC attorney, for his close attention to the paper and helpful suggestions for making certain points more precise. That said, any errors in this paper are solely those of the author.

2 *E.g., FTC v. Invention Submission Corp.*, 965 F.2d 1086, 1089-90 (D.C. Cir. 1992).

Third, in contrast to litigation, which is usually triggered by an event resulting in one party having a claim against another, government investigations may be triggered by anything that piques the government's curiosity. They may be triggered by news stories, consumer complaints, requests from Congress, leaks from informants, first-hand observations by government employees, self-reporting, or any number of other sources. An agency "can investigate merely on suspicion that the law is being violated, or even just because it wants assurance that it is not."³

In those relatively few instances in which government investigations end up in court (generally because the government has sued to secure compliance), the courts have consistently recognized that the scope of issues that they may consider "must be narrow, because of the important governmental interest in the expeditious investigation of possible unlawful activity."⁴ Although the court's function in such proceedings is "neither minor nor ministerial,"⁵ it is "strictly limited" to determining whether the inquiry is "within the authority of the agency, the demand is not too indefinite and the information sought is reasonably relevant" [and] 'the disclosure sought [is] not [] unreasonable."⁶ Significantly, a government request is not unreasonably burdensome unless it "threatens to unduly disrupt or seriously hinder normal business operations."⁷

B. The Government's Civil Investigative Arsenal

Under Part V of the Federal Rules of Civil Procedure (Rules 26-37) and Rule 45, litigants have several paths for discovering information; and courts may apply a broad range of sanctions to compel, or at least encourage, cooperation. In contrast, the government in its investigations must depend on statutory grants of authority to obtain information. Absent voluntary cooperation or statutory grants of authority, the government is powerless to collect information before filing any legal action.

The Federal Trade Commission ("FTC"), which has a full range of information-gathering resources at its disposal, is a good example of both the breadth and limits of the government's ability to gather information in a pre-complaint investigation. At one end of the spectrum, the agency's statutes allow it to – and in practice it does – encourage voluntary cooperation by issuing access letters,⁸ which are unenforceable requests for information.⁹ In this regard, the FTC Act protects the information from public disclosure. It provides that the agency will afford information given "in place of compulsory . . . process" the same level of confidential treatment that it provides to information it receives through compulsory process.¹⁰

On the other end of the spectrum, the agency may compel parties to give information. For example, the FTC may issue orders directing persons to submit "special reports" providing detailed information about their conduct and other matters.¹¹ Such orders are judicially enforceable,¹² and failure to comply may result in the imposition of

3 *FTC v. Texaco, Inc.*, 555 F.2d 862, 872 (D.C. Cir. 1977) (*en banc*) (quoting *United States v. Morton Salt Co.*, 338 U.S. 632, 642-43 (1950)).

4 *Texaco*, 555 F.2d at 872.

5 *Okla. Press Publ'g Co. v. Walling*, 327 U.S. 186, 217 n.57 (1946).

6 *Morton Salt Co.*, 338 U.S. at 652-53 (1950); *Texaco*, 555 F.2d at 872.

7 *Texaco*, 555 F.2d at 882; *Invention Submission Corp.*, 965 F.2d at 1090.

8 15 U.S.C. §§ 46(a), 49; see also FTC Rule 2.4, 16 C.F.R. § 2.4.

9 See *FTC v. Am. Tobacco Co.*, 264 U.S. 298 (1924).

10 See 15 U.S.C. § 57b-2(f); FTC Rule 4.10(a)(8)(ii), 16 C.F.R. § 4.10(a)(8)(ii).

11 15 U.S.C. § 46(b). The Commission's authority to order such reports is limited to investigations that do not involve unfair or deceptive acts or practices. 15 U.S.C. § 57b-1(b).

12 15 U.S.C. § 49.

civil penalties, which accrue daily.¹³ In all its investigations, the agency also has the authority to issue civil investigative demands (“CIDs”) that may compel the recipient to provide information through interrogatory-style questions, produce documentary materials, or appear and give testimony at investigational hearings.¹⁴ In its antitrust investigations, the agency additionally has the power to issue administrative subpoenas to compel the production of documents or the giving of testimony at investigational hearings almost anywhere in the country.¹⁵ FTC CIDs and subpoenas are both judicially enforceable, and those who do not comply with a court’s enforcement order may face contempt charges.¹⁶

Other federal agencies have additional powers to gather information. For example, the Consumer Financial Protection Bureau (“CFPB”) has supervisory authority that provides it with immediate access to certain records maintained by entities it regulates.¹⁷ Similarly, the Securities and Exchange Commission (“SEC”) requires certain entities subject to its jurisdiction to file financial and other reports, and to maintain certain records that the agency may see at any time.¹⁸ That agency also asserts the authority to issue “forthwith subpoenas.”¹⁹ Similarly, the Department of Labor and the Environmental Protection Agency both mandate the retention of various records; and in some instances the failure to maintain the records can result in fines or even imprisonment.²⁰

The premerger notification statute, the Hart-Scott-Rodino Act (“HSR Act”),²¹ lies somewhere between “voluntary” and “compulsory.” On the one hand, the HSR Act authorizes the antitrust agencies to request detailed information relating to covered transactions. On the other hand, the parties are not required to respond to the requests – although they are forbidden to consummate their transaction unless they observe a statutory waiting period after providing either all the requested information or a detailed statement of reasons why they cannot provide the information.²² Failure to comply with the HSR reporting and waiting-period requirement may trigger a court action to enjoin the transaction until there has been compliance,²³ an action to rescind the transaction if it has been consummated,²⁴ or a suit for substantial civil penalties, which accrue daily.²⁵ As these examples show, the government has the tools it needs to conduct its investigations, and it has had those tools for some time.

Access letters and subpoenas have been in the FTC’s toolbox since the very beginning. It gained its CID authority in consumer protection cases in 1980 and in

13 15 U.S.C. § 50.

14 15 U.S.C. § 57b-1.

15 15 U.S.C. § 49.

16 15 U.S.C. §§ 49, 57b-1(h). *See, e.g.*, Stipulation for Entry of Order, *FTC v. Western Union Co.*, No. 13-mc-00131 AKH (S.D.N.Y. Dec. 9, 2013), ECF No. 67. Other agencies have also successfully sought civil contempt against parties who disobeyed court orders. *See, e.g.*, Contempt Order, *SEC v. Coronati*, No. 13-Misc.-372-P1 (S.D.N.Y. Jan. 17, 2014), ECF No. 19.

17 *See* 12 U.S.C. §§ 5512(b)(1), 5514(a)(1)(C); 5514(b)(7); 12 C.F.R. pt. 1091.

18 *See, e.g.*, 15 U.S.C. §§ 78m, 78q; 17 C.F.R. §§ 230.401-498, 240.12b-1 to -37.

19 For a discussion of “forthwith subpoenas,” *see* John Reed Stark, *When to Say When: Handling Emerging Technology-Related SEC Enforcement Tactics*, 45 SEC. REG. & L. REP. (BNA) 1737 (Sept. 23, 2013), available at http://www.strozfriedberg.com/wp-content/uploads/2013/09/When-to-Say-When-Handling-Emerging-Technology-Related-SEC-Enforcement-Tactics_BloombergBNA_Stark.pdf.

20 *See, e.g.*, 29 C.F.R. pt. 516; 40 C.F.R. §§ 141.31-35. For a summary of some Department of Labor and EPA record keeping requirements, *see* *Wages: Record Keeping & Reporting*, U.S. DEPARTMENT OF LABOR, <http://www.dol.gov/dol/topic/wages/wagesrecordkeeping.htm> (last visited July 16, 2014); Office of Water, Env’tl. Prot. Agency, EPA 816-F-06-033, Record Keeping Rules: A Quick Reference Guide (2006), available at http://www.epa.gov/ogwdw/smallsystems/pdfs/guide_smallsystems_records_08-25-06.pdf.

21 15 U.S.C. § 18a.

22 *See* 18 U.S.C. § 18a(e); 16 C.F.R. § 803.3.

23 15 U.S.C. § 18a(g)(2).

24 *See, e.g.*, *FTC v. Elders Grain, Inc.*, 868 F.2d 901 (7th Cir. 1989) (granting rescission on the merits).

25 15 U.S.C. § 18a(g)(1).

competition cases in 1994; and the HSR Act has been in effect since 1978. While the FTC's investigative tools have remained constant since 1994,²⁶ FTC antitrust investigations have grown in size and complexity. In the mid-1990s, very few cases involved document productions exceeding one million pages, and significant major merger investigations might have resulted in the production of a couple hundred boxes of documents.²⁷ Today, FTC merger investigations may yield terabytes of information.

While some may argue that these numbers evince the growth of intrusive government regulation, two facts account for the government's increased demand for information. First, the public, the courts, and the Congress all (correctly) demand that solid public interest justifications underpin any government intrusions into private decision-making. In matters involving private economic activity, this means agencies must base regulatory actions on evidence showing, in one form or another, that the public benefit from regulation is sufficient – or at least probable enough – to offset any private harm that may follow from the regulation. The Supreme Court's 1974 *General Dynamics*²⁸ decision is illustrative. That decision foreshadowed the end of the government's ability to prove a Clayton Act violation with simple evidence showing that a given merger would result in highly concentrated markets. While the government formerly could prove a violation in merger cases by simply showing an undue increase in four- or eight-firm concentration ratios,²⁹ or even that a very large firm was acquiring a very small one,³⁰ today the government must produce solid economic evidence showing that a merger may substantially lessen competition if consummated.³¹ This often requires sophisticated economic analysis and modeling. This evidentiary burden requires the government to collect substantial data and information from the merging parties and other persons.³²

Second, the quantity of potential evidence is vastly greater today than it was in the past. Virtually everyone with any decision-making authority in today's business environment has an array of electronic devices and social media tools readily at hand and uses them to transact business and communicate with superiors, co-workers, subordinates, and outside parties. Email, voice mail, instant messages, text messages, tweets, word processing, spreadsheets, presentations, and data compilations move freely and quickly through (and outside) an enterprise, and in various forms may be modified, preserved, replicated, and archived in the process. Sometimes employees work around the information structure of the enterprise and carry or transmit data and information off-site. In addition, companies and their employees are increasingly using "cloud"-based providers for hosting and processing data, as well as social networking sites, which are all hosted by third parties. Further, organizations increasingly allow employees to conduct business on any device they choose, rather than restricting them to company-owned equipment – a phenomenon

26 The SAFEWEB Act expanded the Commission's ability to secure and share information in some circumstances with foreign governments, but did not *per se* enlarge the tools at the FTC's disposal. See 15 U.S.C. § 46(j).

27 *Cf.*, Announcement by Deborah Platt Majoras, Chairman, Federal Trade Commission, Reforms to the Merger Review Process 5-6 (Feb. 16, 2006) [hereafter referred to as "Reforms to Merger Review"], available at <http://www.ftc.gov/sites/default/files/attachments/mergers/mergerreviewprocess.pdf>.

28 *United States v. General Dynamics Corp.*, 415 U.S. 486 (1974).

29 *See, e.g., United States v. Philadelphia Nat'l Bank*, 374 U.S. 321 (1963).

30 *See, e.g., United States v. Aluminum Co. of Am.*, 377 U.S. 271, 278-81 (1964).

31 *See, e.g., FTC v. Whole Foods Market, Inc.*, 548 F.3d 1028 (D.C. Cir. 2008); *FTC v. H.J. Heinz Co.*, 246 F.3d 708, 716-18 (D.C. Cir. 2001).

32 *See, e.g., FTC v. Staples, Inc.*, 970 F. Supp. 1066 (D.D.C. 1997). The FTC's lawsuit against Blockbuster further illustrates this point. In March, 2005, the FTC sued to enjoin Blockbuster's acquisition of Hollywood Entertainment Corp., on the ground that it had not complied with the premerger notification reporting and waiting period requirements. The Commission's complaint alleged, among other things, that Blockbuster had provided data for only approximately 400 of the company's 4,600 stores and that, at the time of the complaint, Blockbuster had only recently corrected the problem. The Commission alleged, "[t]he original data disk produced by Blockbuster contained 2.8 megabytes of data and had approximately 65,000 data rows [while the corrected disk] contained 96 megabytes of data and approximately 873,000 data rows." Complaint at § 17, *FTC v. Blockbuster, Inc.*, No. 1:05CV00463 (D.D.C. Mar. 4, 2005), ECF No. 1, available at <http://www.ftc.gov/sites/default/files/documents/cases/2005/03/050304compblockbuster.pdf>.

known as “bring your own device.” Thus, important information can be widely dispersed through corporate networks, third-party servers, company- and privately-owned devices, electronic media storage (such as thumb drives and CD-ROMS), and Internet websites.

In the not-too-distant past, those engaging in questionable acts could hope to avoid detection by the simple expedient of circulating undated, unsigned memos with no letterhead. Even if discovered, ownership, distribution, and authorship of such documents was easy to deny, or at least not recall. Today’s digital world makes such evasion all but impossible – provided government investigators get their hands on the right devices or sources. Of course, in order to retrieve all the relevant electronic evidence and identify those with knowledge of it, the government must cast a broad net; and this need often results in substantial demands on investigative targets and others.

C. How Government Investigations Begin

Parties typically have little opportunity to shape a government investigation at its very earliest stages. Information that triggers an investigation may come from many different sources, ranging from news reports to consumer complaints, leaked information, Congressional inquiries, or even reports or information that parties themselves file with the government. Just as the sources that trigger an investigation vary widely, so do the procedures that various agencies follow in opening an investigation. For example, in some agencies, such as the EPA, CFTC, and SEC, investigators are delegated broad discretion in choosing whether to follow a “hot tip.” In contrast, agencies such as the FTC exercise top-down control over the process, while others, such as the CFPB, fall in between.³³

Statutory requirements shape the process used at the FTC. Accordingly, FTC staff may conduct only a limited inquiry to see whether an alleged act or practice warrants closer examination. Assuming staff thinks a matter is worth investigating, the agency’s attorneys must prepare a written recommendation that the Commission open a full investigation and authorize staff to use compulsory process. These “process memos” are not cursory. Staff must identify the target, or potential targets, of the inquiry, the conduct that is of concern, the ways in which that conduct may violate any law that the Commission enforces, the sort of evidence the staff believes it will need to collect to determine whether there may be a law violation, the possible legal and factual defenses the targets may use to counter any legal challenge, and how staff plans to proceed with the investigation. The Commission then opens a formal investigation only if a majority of voting Commissioners approve it, at which time the Commission will issue a Resolution Authorizing Compulsory Process. The Resolution identifies the target or potential targets of the investigation, the conduct that is being investigated, and the legal basis for the inquiry.

Notably, the Resolution is not the Commission’s last contact with the matter. Although the staff is responsible for drawing up the specifications for each subpoena and CID, they have no authority to actually issue compulsory process. By law, all compulsory process issued by the Commission must be signed by an individual Commissioner. This means that staff must prepare the papers and submit them to the Commissioner assigned to the matter, who in turn must review and sign them before they may be served by the Commission’s Secretary.

³³ The CFPB’s Rules of Practice specify that only an Assistant Director or Deputy Assistant Director of the Office of Enforcement has authority to initiate investigations and issue process. 12 C.F.R. §§ 1080.4, 1080.6(a). The Rules further limit the authority to close investigations to the Assistant Director or Deputy Assistant Director. *Id.* § 1080.11.

No matter how government investigations may begin, civil law enforcement investigations typically fall into two categories: those that the parties cannot anticipate because they are not aware of the government's concerns about a matter, and those that they can anticipate because the parties are engaging in activities that are likely to trigger an inquiry. In the former situation, the parties usually have little or no opportunity to shape the government's inquiry because the groundwork for the inquiry is laid before the investigation officially begins. However, parties who can anticipate an investigation and who are willing to engage the government proactively – sometimes even before any event can trigger an inquiry – do have an opportunity to shape the investigation, by discussing matters that they believe may raise particular concerns. Parties who are candid and cooperative in this early engagement have a unique potential to focus and limit the scope of an inquiry. Indeed, some practitioners boast of their “track record” in “working things out” before agencies even open inquiries.

D. Options for Responding to Government Civil Investigations

Targets of government investigations seem to employ one of three methods in responding. First, some resist by delaying every response, seemingly nitpicking over every document request, construing every request narrowly, and litigating – or threatening court challenges – at every opportunity. Second, some take an arm's-length approach. They volunteer nothing, leave the government to figure out what it needs, and surrender only what is requested when threatened with enforcement. They engage in dialogue with the government only when, and if, the government starts the conversation. Third, others cooperate by engaging in early and frequent discussions with investigators to determine what the government needs, providing the requested materials on time, and proactively working with the government to find the best way to address its concerns.

The difficulty with the first two approaches is that it is impossible for the government investigator, who is trained to identify suspicious activities, to distinguish between those who have something to hide, those who have nothing to hide but are clueless about the process, and those who are merely taking a “make-them-work-for-it” approach. Admittedly, practitioners who take either of the first two approaches to civil investigations may well be seeking to protect privileged or legitimate but highly confidential business material, or simply trying to advocate their strongly held view of the merits from the outset. However, the approach entails a big and risky bet that government investigators will back off, either from exhaustion or intimidation, and not pursue an investigation thoroughly if the target plays hardball. On balance, this seems like placing a substantial bet that could cost the client dearly in the long run, as the client consumes human and capital resources while the investigation methodically grinds through each new lead the investigators uncover.³⁴ As noted in Parts A and B, *supra*, the government generally has the means to obtain the information it needs.

The third option, that of full cooperation, does not preclude a practitioner from maintaining an arm's-length relationship with the government and fully and vigorously representing a client's interests. Cooperation does not require social interactions, but it does require honesty and candor. Admittedly, the approach all but guarantees that the government will find any relevant information and deal with any law violation that it

³⁴ Note too, that, while courts sometimes require the requesting party in litigation to bear part of the cost of discovery (Fed. R. Civ. P. 26(b)(4)(C)), cost-shifting mechanisms are rarely available to civil investigative targets, or even to third parties. *See, e.g., Texaco*, 555 F.2d at 882 (enforcement of compulsory process will not be denied on grounds of burdensomeness and breadth, absent a showing that “compliance threatens to unduly disrupt or seriously hinder normal operations of a business”), *cert. denied*, 431 U.S. 914 (1977).

uncovers. Nonetheless, cooperation offers several distinct advantages to the target. Even when it cannot lead to leniency, it at least gives the target an opportunity to focus and narrow the government's inquiry, with a consequent speedy and relatively less expensive resolution of the matter.³⁵ The balance of this paper identifies some of the things a target can do to ensure such an outcome.

E. Practical Means of Narrowing and Limiting Law Enforcement Investigations

Two overarching facts greatly influence the course of many law enforcement investigations. They are these:

First:

The government typically does not know the organizational structure of any specific corporation, or the manner in which it creates, distributes, analyzes, uses, retains, and destroys records.

Second:

Government investigators are often required to complete investigations promptly, often within rigid deadlines, and without making repeated demands for information.

Investigators must keep these facts in mind when shaping their requests for documents and information. Therefore, their instructions regarding the definition of the "target," and the scope of the expected search for responsive information, may look something like this:

The corporation includes its domestic and foreign parents, predecessors, divisions, subsidiaries, affiliates, partnerships and joint ventures, and all directors, officers, employees, agents and representatives of the foregoing. The terms "subsidiary," "affiliate" and "joint venture" refer to any person in which there is partial (25 percent or more) or total ownership or control between the company and any other person.³⁶

In short, the instructions tell a party that to comply fully it must search every desk, person and file drawer, even in its affiliates' offices, as well as every computer, server, cloud-based source, notebook, smartphone, phone mail system, tablet, and other device that may hold responsive information.

The instructions for producing computer files are similarly comprehensive, specifying whether files must be submitted in native format, in image format with extracted text and metadata, image format accompanied by OCR, or some combination thereof. The instructions also identify the metadata fields that must be submitted, the use of de-

³⁵ In some circumstances, cooperation can even lead to less burdensome remedial orders, or even no order at all. *See* Sec. & Exch. Comm'n, Accounting and Auditing Enforcement Release No. 1470, Report of Investigation Pursuant to Section 21(a) of the Securities Exchange Act of 1934 and Commission Statement on the Relationship to Agency Enforcement Decisions (Oct. 23, 2001), *available at* <http://www.sec.gov/litigation/investreport/34-44969.htm>.

³⁶ *See* Premerger Notification Office, Fed. Trade Comm'n, Hart-Scott-Rodino Premerger Notification Program Introductory Guide III: Model Request for Additional Information and Documentary Material (Second Request) 10 (June 2010), *available at* <http://www.ftc.gov/sites/default/files/attachments/premerger-introductory-guides/guide3.pdf>.

duplication or email threading software, the criteria for submitting data, such as data in Excel spreadsheets, and the media to be used in submitted productions.³⁷

Unless the recipient of an investigative demand is prepared to face the potential consequences of conducting an inadequate search or of having important evidence obliterated, its lawyers should immediately talk to the investigators. That conversation should address, at minimum, the following six subjects: (1) the scope of the search, both as to time and as to custodians; (2) data preservation and retrieval issues, including email, phone mail, tweets and instant messages, social media, and cloud sources; (3) the retention and disposition of legacy systems, archives, and backup tapes; (4) privilege logs; (5) materials located outside the United States; and (6) the timing and staging of production.

1. Implementing a Litigation Hold / Directing Preservation

The first step that a party must take upon learning of an investigation is to implement a litigation hold. At common law, the duty to preserve evidence attaches when a person with possession, custody, or control over the evidence reasonably anticipates litigation in which that evidence may be relevant. Under Fed. R. Civ. P. 37, a party who fails to produce relevant evidence may face substantial sanctions including the entry of a default judgment against the wrongdoer or the imposition of substantial costs. The government, during an investigation, does not, of course, have Rule 37 sanctions available to it – at least not until after it files a case – but, as discussed in Part A, *supra*, it does have means of enforcing its pre-complaint discovery demands and, in some instances, of extracting penalties for non-compliance. Laying aside those instances in which fines and penalties are more or less automatic,³⁸ the most significant tools are the ability to seek civil contempt against parties who do not obey court orders enforcing process,³⁹ and the ability to charge obstruction as a crime.⁴⁰

The potential for civil and criminal liability shows the importance of parties taking immediate steps to preserve information and materials when they have notice of an actual or contemplated government inquiry. Even when the party does not expect a government inquiry to lead to litigation, the party must preserve relevant information. This is so because the government has the right to conduct investigations, even if it is only “seeking assurance” that the law is not being violated. Also, in many situations, it has the authority to “investigate” matters, for the purpose of preparing a study or a report.⁴¹ In short, law suits and sanctions are not always the object of investigations. Parties who ignore or, worse yet, “dispose” of information responsive to even an “informal” inquiry do so at their peril.

The principles that govern retention in investigations are the same principles that govern retention in the civil litigation: parties are to take prompt and reasonable, not herculean, steps to preserve and to stop the routine destruction and disposition of relevant materials. This requires identifying the custodians of relevant information, informing them of their obligation to preserve relevant information and materials, and following up to ensure that they are complying. It also means identifying all other sources of potentially relevant electronic information and implementing procedures to ensure that such information will not be destroyed.

³⁷ *Id.* at 13-14.

³⁸ For example, Section 21(c) of the Securities Exchange Act, 15 U.S.C. § 78u(c), potentially makes it a misdemeanor, punishable by fines and imprisonment, to wilfully disobey an SEC subpoena.

³⁹ Indeed, civil contempt can sometimes result in imprisonment. *See, e.g., Coronati, supra* note 16.

⁴⁰ 18 U.S.C. §§ 1505, 1506, 1509, 1519, 1520.

⁴¹ *E.g., Morton Salt Co.*, 338 U.S. 632.

2. Preparation: Assessing the Landscape / Developing A Plan

The second step in dealing with a government inquiry is to develop a realistic discovery and disclosure plan. In general, files belong to one of two groups. Either they are “corporate” files found in centralized storage places (such as shared network folders or workspaces, databases, or cloud-based storage applications), or they are “custodian-based” files that are either (1) network-based files that are readily associated with particular custodians (such as email or voicemail) or (2) files maintained off-network in localized sources (such as file drawers, personal computers, smartphones or portable media (tablets and flash drives)). These groups, and the type and volume of accessible data within each group, shape negotiations about the scope of the search and ultimate production.

The subject matter of the investigation and the period it covers define the boundaries of the search. Accordingly, the way to limit the search is to talk to the investigators about the issues that are of concern and the periods and sources of relevant information that will need to be produced. To do this effectively, counsel need to know the structure and content of corporate data and any special costs or burdens associated with retrieving those files. Once counsel defines and understands the scope of what may be relevant, he or she can draw distinctions between the corporate sources that are essential, those that are marginal and might not need to be searched or reviewed if the essential files are sufficient to satisfy the government’s needs, and those that are irrelevant or only arguably relevant.

Custodian-based files present a different problem. Regardless of the issues, the ultimate production of such data requires a search of each custodian’s files to separate the relevant from the irrelevant and the privileged relevant from the non-privileged relevant. Even with the use of technology-assisted review and other efficiency-enhancing technologies and processes, this is a very expensive and labor-intensive effort. Thus, one important key to minimizing cost and burden is to limit the number of custodians. This requires a thorough understanding of both the formal and informal organizational structure, the allocation of job responsibilities, and the way in which people and offices communicate and interact. With that knowledge, counsel can identify the personnel who have direct knowledge of the relevant issues, those who have no knowledge (even though to an outsider it may appear otherwise), and those who are somewhere in the middle. After making these determinations, counsel may propose a list of custodians and undertake some sample searches to show how the search plan will work. He or she might also be able to propose a staged production, where key custodian files are produced first and files from other “relevant but not critical” custodians are deferred.

Counsel also must bear in mind that some electronic files most likely must be produced in native form. As the Federal Trade Commission has noted:

[P]rinted versions of Microsoft Excel spreadsheets are inherently inadequate, because they do not include cell contents, comments, and formulas. Similarly, many programs generate conflicts when their files are printed on popular printers; such conflicts may, for example, eliminate or change underlined or bolded characters, or result in the failure to show the existence of attachments. Further, electronic documents contain “metadata” – embedded data that does not print with the document, but

which includes vital information such as bibliographic data about the document and the names of the recipients of “blind” copies on emails.⁴²

Thus, counsel must first determine which data sets are relevant to the investigation, then learn how the client collects, maintains, and uses that data as well as any software used to maintain and analyze it. In developing the response plan, counsel should consider preparing data samples to demonstrate the types of information the target maintains and the capabilities of its systems and software to sort and analyze electronic information.

Email and other digital communications are the primary means of communication today. They also enable wide input into final written materials. As a result, digital messages, their attachments, and draft documents all have an uncanny ability to show up in unexpected times and places. They can be expensive to deal with because of their enormous volume. If an investigator's demand for them is not limited, they must all be collected, processed, analyzed and reviewed for responsiveness and privilege, and if privileged, logged.

Besides limiting the number of custodians whose files must be searched, a second method – one that works not only for email and word processing documents, but for other electronic information as well – is to use one or more advanced technology options that utilize computer software to assist in determining “relevance” based on user-selected criteria, or “seed sets.” These applications, sometimes referred to as “technology-assisted review,” may be extremely effective and efficient when properly used and verified. Other techniques, such as the use of search terms, concept clustering, de-duplication, near duplication, and email threading, can also yield efficiencies. Each variation of each technology is different, and the results are greatly influenced by the human process used to employ the technology. Thus, early in the planning stage counsel should design his or her project management workflow, including which specific technology applications will be used. Not all government investigators may want to know the details of the target's processes and technologies, but targets are increasingly finding themselves being asked such questions. For example, the Department of Justice Antitrust Division usually requires a fulsome disclosure about methodology before agreeing to a party's use of technology assisted review. In contrast, the FTC tends to seek such information much later in the investigation, and only if it thinks a production has been deficient. Ultimately, in presenting any technological approach to government investigators, candor and transparency will be critical to acceptance of the final production.

Next, counsel should assess the periods that may be relevant for each set of relevant documents and data. A reduction in the period that must be searched can result in a substantial savings. For example, the FTC reports that in merger investigations that sought documents for a three-year period, approximately 25% of the documents produced were more than two years old.⁴³ Obviously, a reduction in the period covered by an investigation can result in substantial savings in search, review, and production costs. However, counsel should be mindful that a one-size-fits-all approach may not be appropriate for all searches, even within the same investigation. For example, some sales data may make sense only when viewed over several seasonal cycles. Similarly, it may be appropriate to take a longer look back into the individual emails of some employees, while a shorter period may be appropriate for others. Counsel should draw rational lines when seeking to limit discovery periods.

42 Fed. Trade Comm'n, Statement of the Federal Trade Commission's Bureau of Competition On Guidelines for Merger Investigations 4 [hereafter referred to as “Bureau of Competition Guidelines”], available at, http://www.ftc.gov/system/files/documents/public_events/114015/ftc_statement_on_guidelines_for_merger_investigations_12-22-02_2.pdf.

43 Reforms to Merger Review, *supra* note 27, at 19.

After counsel has identified the appropriate files and custodians, and the appropriate periods for searching each, he or she can develop a systematic plan and methodology for searching and producing relevant information from those files.⁴⁴ Depending on the case and circumstance, that plan might include a suggestion that the investigation proceed in a layered fashion, whereby the target first produces “core” files, and the government agrees to give those files at least a preliminary look before determining whether to require additional information.⁴⁵

3. Presenting the Plan

The third step in dealing with the inquiry is to convince the investigators to accept the discovery plan, or something close to it. Here it is important to know the Rules of Practice of the agency conducting the investigation. While all agencies encourage parties to meet and confer with investigators, some require it.⁴⁶ Similarly, the FTC’s Reforms To Merger Review; Bureau of Competition Guidelines; and its Bureau of Economics’ Best Practices for Data, and Economics and Financial Analyses in Antitrust Investigations⁴⁷ identify several steps that counsel may take to streamline and facilitate complex investigations. These steps may be synthesized as follows:

- Meet with the investigators as soon as possible. When the parties anticipate an investigation, this may mean meeting before the investigation is formally opened.
- Provide the investigators with organization charts or equivalent materials so they can identify the parties’ employees and their positions.
- Provide the investigators with brief written descriptions of the responsibilities of each person the investigators identify as a person whose files might be searched.
- Present the discovery plan and ideally provide sample search results so the investigators can assess the plan and methodology.
- Make one or more knowledgeable people readily (and repeatedly) available to the investigators. These people should be knowledgeable about the issues and be able to assist the investigators in identifying people whose files must be searched.
- Discuss with the investigators the types and forms of electronic data the parties maintain and provide data samples to assist them in determining what data and data compilations are available.

⁴⁴ The mechanics of document preservation instructions and litigation holds are beyond the scope of this paper; but implicit in this paper is the assumption that a party will issue appropriate instructions to its employees as soon as it identifies files that may be relevant or responsive to the inquiry.

⁴⁵ Except in the most complex mergers, the FTC’s policy in premerger investigations is to limit to 35 the number of persons whose files must be searched. Notably, if a person is in the “search group,” the search must extend to those who maintain his or her files as well as that person’s “personal assistants, secretary, or person with the same or similar responsibilities.” Also, that limitation is subject to receiving full cooperation from the merging party in identifying the appropriate files to search. *See* Reforms to Merger Review, *supra* note 27, at 9-11.

⁴⁶ *See* FTC Rule 2.7(k), 16 C.F.R. 2.7(k). In contrast to the FTC, the SEC’s Rules neither require a meet and confer nor provide a means to ask the agency to quash or modify an administrative subpoena. Nonetheless, that agency encourages parties to meet with staff and discuss any and all issues relating to compliance.

⁴⁷ *Economics Best Practices*, FEDERAL TRADE COMMISSION, <http://www.ftc.gov/about-ftc/bureaus-offices/bureau-economics/best-practices> (last visited July 16, 2014).

- Make available to the investigators one or more people thoroughly knowledgeable about the parties' computer systems and software and the way in which the parties collect, store, maintain, analyze, and use the data and other electronic information that is relevant to the investigation.
- Where appropriate to the investigation, discuss the parties' own economic or financial analyses with (and suggest appropriate analyses to) the investigators. In doing so, the parties should provide backup data and information to enable the investigators to test the parties' data, programs, and results.
- Consider submitting "white papers" that address the issues and provide a sound analysis of the issues from the parties' perspective.

The goal is to provide the government with the relevant information it needs to finish its investigation, while limiting the cost and burden for the target, particularly with respect to the production of information that may be unnecessary, duplicative, or only tangentially relevant. Phased or prioritized discovery can often achieve these goals, particularly when combined with a good faith effort to address the government's concerns through voluntary submissions on the merits.

4. Privilege Issues

Counsel for parties should discuss two privilege-related issues with the investigators: waivers and privilege logs. In some situations, parties are willing to knowingly waive privilege claims and allow investigators to review at least some of their privileged materials.⁴⁸ At the same time, some agencies, including the FTC, have policies or practices of returning privileged documents that are truly produced unintentionally.⁴⁹ If parties intend to waive any privilege claims, they should make this clear at the outset so the investigators are able to distinguish between the documents they may review and those they must set aside to determine if they should be returned.

Complete privilege logs can be time-consuming and expensive to produce. Yet, the information in such logs is essential to investigators who need to determine whether documents are being withheld improperly. An agreement concerning the preparation of a partial log can save time and money while meeting the needs of the investigators. For example, a party may suggest submitting a partial privilege log in which it merely identifies each person who has custody of documents claimed to be privileged and the number of documents each such person holds. In response, the investigators might then designate a smaller subset of custodians whose files must be fully logged. Of course, in this scenario the agency would reserve the right to demand a full privilege log should the matter proceed to litigation.⁵⁰

48 This paper does not discuss the scope or implications of such intentional waivers and does not consider, at all, whether they result in waivers as to third parties.

49 See, e.g., Commission Rule 2.11(d)(1)(ii), 16 C.F.R. § 2.11(d)(1)(ii). Similarly, in the merger context the Commission has said: "By 'inadvertent production' [the FTC] refer[s] to the established body of case law that defines truly inadvertent production as a mistake that occurs despite the existence and use of reasonable procedures to screen out privileged materials. This situation differs from production that occurs because of negligence so significant that – taking into account the totality of the circumstances, including the extent and timing of production – it may still constitute a waiver." Bureau of Competition Guidelines, *supra* note 42, at 3 (footnote omitted).

50 Reforms to Merger Review, *supra* note 27, at 25-26. In this regard, the FTC's Rule 2.11(b) makes clear the agency's willingness to explore meaningful alternatives to the production of full privilege logs. See 16 C.F.R. § 2.11(b).

With respect to both waiver and the potential for reducing the privilege review/logging burden, civil litigants can rely on the protections of Federal Rule of Evidence 502. That Rule governs the disclosure of privileged information in court proceedings or “to a federal office or agency,” and potentially offers some prospect for relief from detailed privilege review in the context of law enforcement investigations. In brief, the Rule applies to work product materials and attorney-client communications. It provides that the voluntary disclosure of such information usually results only in a waiver of the information disclosed (Rule 502(a)) and the involuntary disclosure of such information results in no waiver (Rule 502(b)), if certain criteria are met. The Rule further provides that agreements relating to the disclosure of privileged information (often referred to as “claw back agreements”) are binding only on the parties to the agreement (Rule 502(e)); but such agreements will bind non-parties if the agreements are incorporated into court orders (Rule 502(d)).

Although Rule 502 leaves a gap with respect to pre-litigation agreements, such as those reached in a government investigation, Rule 502(d) and (e) suggest there is room for the government and private parties to negotiate claw back or quick peek agreements to facilitate privilege review during investigations. Although such agreements would necessarily be reached before any court complaint is filed, courts in any subsequent proceedings – either in law enforcement actions or in unrelated actions seeking access to the information provided to the government – would do well to give effect to the purpose of Rule 502 and hold that such agreements do not constitute subject matter waivers. Alternatively, in appropriate cases the government might file a subpoena enforcement action and simultaneously ask the court to “So Order” a settlement that incorporates a claw back or quick peek agreement under its Rule 502(d) authority. Here too, the FTC’s Rule 2.11(d) closely tracks Rule 502 and signals the agency’s willingness to work with parties on these issues.

5. Legacy Systems, Archives, and Backup Tapes

If the government suspects a law violation, and its investigators believe that some electronic information has been recently deleted, it will be keenly interested in obtaining information from alternative sources. As the FTC’s Bureau of Competition has stated, “in our experience, in some cases the search of even a small portion of the parties’ archive and backup systems produces valuable information that is helpful to the staff’s investigation.”⁵¹ However, the FTC also recognizes that backup tapes are not always configured for routine document collection when they are intended solely for disaster recovery or archiving purposes and that review of backup tapes “is expensive and may be duplicative.”⁵²

To balance the potential cost to private parties of reviewing disaster recovery tapes or other non-accessible sources against the potential benefit to the government (and the public) in securing relevant evidence, the FTC’s policy in merger cases is to “require a party to produce documents contained on backup tapes only when responsive documents are not available through other more accessible sources.”⁵³ However, if a party uses backup tapes as its sole means of preserving material subject to a litigation hold or relevant to the investigation, it should expect the FTC to demand that the backup tapes be searched for relevant information. The question, at least initially, is not whether such material must be produced. Rather, the question is whether and how such information must be preserved, pending a determination that information, or some subset thereof, must be searched.

51 Bureau of Competition Guidelines, *supra* note 42, at § 6(c).

52 *Id.*

53 Reforms to Merger Review, *supra* note 27, at 24.

Here too the FTC's merger review policy statement offers a solution that might be applied in other civil investigations:

[A] party may elect to preserve backup tapes for only two calendar days identified by staff, and . . . [i]f a party's document storage system does not permit designation of backup tapes for two specific calendar days, staff will work with the party to designate a comparable set of backup tapes that the party must preserve.⁵⁴

Investigators might not demand that a party preserve all backup tapes, but only a small subset, which may need to be reviewed in the event the staff determines there are significant gaps in the materials obtained from other sources. However, a party may not unilaterally decide which backup tapes to preserve and which to recycle. That determination is for the agency to make after the party and the agency investigators have met "to discuss information about the archives and backup systems."⁵⁵ This is yet another topic for the dialogue between the government and the counsel, and there can be significant benefits in resolving questions about the preservation of backup tapes and other inaccessible sources, particularly with respect to reducing costs and future litigation risks.

6. Parallel Investigations and International Matters

Private conduct will sometimes interest more than one federal agency, may raise concerns with various state agencies, and will frequently get the attention of foreign authorities as well. In other words, multiple law enforcement agencies and jurisdictions may conduct parallel investigations. Not surprisingly, law enforcement agencies increasingly recognize the advantages of cooperating with one another.⁵⁶ Such cooperation may take the form of sharing information with other federal agencies, granting states access to various federal files, and agreements and memoranda of understanding between the various agencies and foreign law enforcement authorities.⁵⁷ This cooperation has the potential to benefit everyone. On the one hand, the agencies have "an interest in reaching, insofar as possible, consistent, or at least non-conflicting, outcomes."⁵⁸ On the other hand, the parties benefit from speedier resolution of all matters; reduced discovery costs resulting from agency sharing; and less risk of facing conflicting (i.e., mutually exclusive) regulatory requirements.

Notably, the benefits of international cooperation depend largely on the willingness of the investigative target to cooperate in the investigation. Such cooperation may include the granting of waivers to allow the jurisdictions to share information they might otherwise be barred from sharing.⁵⁹ It may also require the parties to engage in multilateral negotiations to coordinate the production of responsive materials and synchronize the investigations so all jurisdictions conclude their investigations at more or less the same time.

⁵⁴ *Id.*

⁵⁵ Bureau of Competition Guidelines, *supra* note 42, at § 6(c).

⁵⁶ See, e.g., U.S. Dept. of Justice & Fed. Trade Comm'n, Antitrust Enforcement Guidelines for International Operations §§ 2.91, 2.92 (April 1995), available at <http://www.usdoj.gov/atr/public/guidelines/international.htm>.

⁵⁷ See, e.g., *International Competition and Consumer Protection Cooperation Agreements*, U.S. DEPARTMENT OF JUSTICE, <http://www.ftc.gov/policy/international/international-cooperation-agreements> (last visited July 16, 2014). An interesting example of international cooperation is the US SAFEWEB Act, which allows the FTC to assist non-US law enforcement agencies under certain conditions. See *In re FTC*, 13-mc-524-MJG, 2014 U.S. Dist. LEXIS 106214 (D. Md. Aug. 4, 2014) (enforcing an FTC Application under 28 U.S.C. § 1782 to obtain information on behalf of the Canadian Competition Bureau).

⁵⁸ US-EU Merger Working Group, Best Practices on Cooperation in Merger Investigations § 1 (Oct. 2011), available at http://www.ftc.gov/system/files/documents/public_statements/310481/111014eumerger.pdf.

⁵⁹ *Id.* §§ 3-7. In this regard, the antitrust agencies have jointly published a Model Waiver of Confidentiality for international civil matters, along with a set of Frequently Asked Questions, to assist parties in determining how and when to waive confidentiality with the enforcement agencies. See *International Waivers of Confidentiality in FTC Antitrust Investigations*, FEDERAL TRADE COMMISSION, <http://ftc.gov/policy/international/international-competition/international-waivers> (last visited July 16, 2014).

CONCLUSION

Because investigators approach each matter on a case-by-case basis, there are no hard and fast rules to inform counsel on which step or combination of steps will succeed in any particular investigation. Nonetheless, the government is generally not anxious to spend time and resources reviewing irrelevant documents and data compilations. It is a rare civil investigation in which the government absolutely must have unlimited access to all the materials conceivably responsive to its original requests. Even in those cases, it is generally willing to talk meaningfully with parties who demonstrate candor and honesty. If a party knows that its conduct will likely result in an order to take corrective action, its best course is likely to “come clean,” get all the facts out, and resolve the issue as quickly as possible. Conversely, if it honestly thinks the government’s investigation is misdirected and unnecessary, the best way to address that is to lay out the facts and let the government satisfy itself that the investigation can be closed. In either circumstance, cooperation will yield a faster, less expensive result than engaging in pitched battles or taking a “make-them-work-for-it” approach.

The key to successfully navigating a client through a government civil investigation lies in understanding the government’s law enforcement concerns and objectives; devising a comprehensive plan for conveying necessary information to the government; and then meeting with the investigators early and frequently throughout the process. Candor and transparency will hasten the process and minimize costs.

WHAT'S THE PROBLEM WITH GOOGLE?

Daniel R. Shulman
Gray, Plant, Mooty, Mooty & Bennett
Minneapolis, MN

I. INTRODUCTION

The antitrust winds have been blowing at gale force around Google, most notably in the EU. This paper will look at the course of the EC proceedings, from the charges to the pending probable settlement. It will then consider the question whether the conduct under discussion can reasonably be found to violate Section 2 of the Sherman Act or abuse of a dominant position under EC law. That it has been so labeled raises some important and fundamental questions of antitrust policy, and suggests a re-examination of the purposes of international competition and trade regulation law.

Google is not only a major international commercial force, but also an extremely beneficial one in many ways. Its success, size, and dominance have also subjected it to heightened antitrust scrutiny. For outsiders, who use Google as the home page on their web browsers and find it an invaluable search engine, the questions naturally arise, "What are the antitrust issues with Google," and "Why do they matter"? This paper will examine these questions, in particular the intersection of antitrust and public policy in areas not usually associated with competition.

II. THE EC PROCEEDINGS

A. Opening the Investigation

On November 30, 2010, the European Commission issued IP/10/1624: "Antitrust: Commission probes allegations of antitrust violations by Google." According to the press release:

The European Commission has decided to open an antitrust investigation into allegations that Google Inc. has abused a dominant position in online search, in violation of European Union rules (Article 102 TFEU). The opening of formal proceedings follows complaints by search service providers about unfavorable treatment of their services in Google's unpaid and sponsored search results coupled with an alleged preferential placement of Google's own services. This initiation of proceedings does not imply that the Commission has proof of any infringements. It only signifies that the Commission will conduct an in-depth investigation of the case as a matter of priority.

The release noted “two types of results when people are searching for information” on Google: “unpaid search results, which are sometimes also referred to as ‘natural’, ‘organic’ or ‘algorithmic’ search results, and third party advertisements shown at the top and at the right hand side of Google’s search results page (so-called paid search results or sponsored links).”

The alleged problem complained of by other “internet search providers” and focused on by the EC was “lowering the ranking of unpaid search results of competing services,” a number of which provided pricing comparisons, and “according preferential placement to the results of its own vertical search services in order to shut out competing services.” The EC also stated that it would look at whether “Google lowered the ‘Quality Score’ for sponsored links of competing vertical search services,” and thereby adversely affected “the price paid to Google by advertisers.” Finally, the EC announced that it would examine “allegations that Google imposes exclusivity obligations on advertising partners, preventing them from placing certain types of competing ads on their web sites, as well as on computer and software vendors, with the aim of shutting out competing search tools,” and whether Google imposed “restrictions on the portability of online advertising campaign data to competing online advertising platforms.”

The nature and scope of the announced proceedings give rise to several observations.

First, the announced impetus for the investigation was competitor complaints, which suggest sour grapes, questionable antitrust injury and standing, and misuse of the antitrust laws to suppress competition and punish a successful rival, rather promoting and preserving competition. *Cargill, Inc. v. Monfort of Colo., Inc.*, 479 U.S. 104, 116 (1984) (“...the antitrust laws do not require the courts to protect small businesses from the loss of profits due to continued competition, but only against the loss of profits from practices forbidden by the antitrust laws.”); *Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc.*, 429 U. S. 477 (1977).

Second, Google’s alleged discrimination against rival search engines and certain advertisers – exclusive dealing requirements excepted – would appear to be behavior condoned even for a monopolist under United States Supreme Court antitrust jurisprudence. The offense of monopolization under Section 2 of the Sherman Act consists of the wilful acquisition or maintenance of monopoly power. *United States v. Grinnell Corp.*, 384 U.S. 563, 570-71 (1966). Monopoly power is “the power to control prices or exclude competition.” *Id.*, at 571. No claim exists that Google obtained monopoly power other than “as a consequence of a superior product, business acumen, or historic accident,” which does not violate Section 2. *Id.*; *United States v. Aluminum Co. of America*, 148 F.2d 416, 430 (2d Cir. 1945; L. Hand, J.) (no violation where a monopoly results from “superior skill, foresight and industry.”).

Hence, the focus must be on whether Google’s discrimination against rival search engines and advertisers constitutes the wilful maintenance of monopoly power. This would consist of “use of that power by anticompetitive or exclusionary means or for anticompetitive or exclusionary purposes.” *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*, 472 US 585, 595-96 (1985). As the Supreme Court further explained, “If a firm has been ‘attempting to exclude rivals on some basis other than efficiency,’ it is fair to characterize its behavior as predatory.” *Id.*, at 605.

Recent Supreme Court decisions, however, have held that a monopolist's choice of its customers and the terms on which it will deal with them are generally not violative of Section 2 except in very limited circumstances. *Communications Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004); *Pacific Bell Telephone Co. v. linkLine Communications, Inc.*, 555 U.S. 438, 129 S.Ct. 1109 (2009).

In *Trinko*, Justice Scalia, writing for the Court, blessed monopoly pricing:

The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices – at least for a short period – is what attracts “business acumen” in the first place; it induces risk taking that produces innovation and economic growth. To safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive *conduct*. [540 U.S. at 407.]

In words that now seem prescient in their application to Google, Justice Scalia continued,

Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities. Enforced sharing also requires antitrust courts to act as central planners, identifying the proper price, quantity, and other terms of dealing – a role for which they are ill-suited. Moreover, compelling negotiation between competitors may facilitate the supreme evil of antitrust: collusion. Thus, as a general matter, the Sherman Act “does not restrict the long recognized right of [a] trader or manufacturer engaged in an entirely private business, freely to exercise his own independent discretion as to parties with whom he will deal.” *United States v. Colgate & Co.*, 250 U. S. 300, 307 (1919). [*Id.* at 407-08.]

In *linkLine*, a unanimous Supreme Court went even further in unshackling monopolists. In an opinion written by Chief Justice Roberts, the Court for the first time recognized an “antitrust duty to deal,” 129 S.Ct. at 1115, which generally is inapplicable to monopolists.

Given the *Trinko* and *linkLine* decisions, and assuming, *arguendo*, convergence in this area of the law with the EC, one must ask what could be amiss with Google giving preferential treatment either to its own search functions or to advertisers and other search firms that are willing to pay the tribute Google requires. Under *Trinko* and *linkLine*, Google had no duty to deal at all with competitors in the search engine business or advertisers that would not meet its demands and prices. The Supreme Court could have said in those decisions, but did not, that if a monopolist chooses to deal with rivals or customers, it must do so on nondiscriminatory and fair terms that do not unduly hamper their ability to compete. Instead, the Court largely left monopolists free to set their own terms and choose their own customers. Hence, the EC accusations against Google appear at first blush to be somewhat surprising.

Third, the EC's pressing ahead in the face of *Trinko* and *linkLine* suggests the possibility of substantial divergence from U.S. antitrust jurisprudence with respect to so-called essential facilities and refusals to deal. In *Trinko*, the Supreme Court cast doubt on the essential facilities doctrine, by refusing to find the plaintiff's claim viable "even if we considered to be established law the 'essential facilities' doctrine crafted by some lower courts." 540 U.S. at 410. In a footnote, the Court distinguished *United States v. Terminal Railroad Assn. of St. Louis*, 224 U. S. 383 (1912), the leading case on the essential facilities doctrine, as involving "concerted action, which presents greater anticompetitive concerns and is amenable to a remedy that does not require judicial estimation of free-market forces: simply requiring that the outsider be granted nondiscriminatory admission to the club."

Prior to *Trinko*, however, lower courts had not so construed *Terminal Railroad*. To the contrary, they had read that case and its progeny as establishing a monopolist's duty to deal under certain specified circumstances:

The case law sets forth four elements necessary to establish liability under the essential facilities doctrine: (1) control of the essential facility by a monopolist; (2) a competitor's inability practically or reasonably to duplicate the essential facility; (3) the denial of the use of the facility to a competitor; and (4) the feasibility of providing the facility.

MCI Communications v. American Tel. & Tel. Co., 708 F.2d 1081, 1132-33 (7th Cir., 1983).¹ The Google investigation by the EC raises the question whether essential facilities law is alive and well in the EU, if not generally, then with respect to the Internet.

Fourth, that the EC is investigating "allegations that Google Inc. has abused a dominant position in online search" implies that it is in fact possible to obtain "a dominant position in online search," a proposition with which not everyone would agree. That bottomless repository of knowledge, Wikipedia, lists, as of the writing of this article, approximately 40 "active" search engines. http://en.wikipedia.org/wiki/Internet_search_engines. The high tech highway is littered with the wrecks of past "dominant" firms (e.g., AOL, BlackBerry, Gateway), while others have had to be reborn from their own ashes of dominance to survive in another form (e.g., IBM, HP), and still others scramble through high-price acquisitions to keep up in order to cling to dominance (Facebook, Apple). Obviously, the question arises whether the Internet is one area where the inexorable working of the marketplace and the irreversible tide of innovation render dominance transitory, if not illusory.

B. The Commission's Preliminary Assessment

On March 13, 2013, the Commission adopted a Preliminary Assessment, in which it found a number of concerns, which it communicated to Google. These included the following "business practices that may violate Article 102" (available at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2013:120:0022:0024:EN:PDF>):

" – the favourable [sic] treatment, within Google's horizontal Web search results, of links to Google's own vertical Web search services as compared to competing vertical Web search services..."

¹ Citing *Hecht v. Pro-Football, Inc.*, 570 F.2d 982, 992-93 (D.C.Cir.1977); *Otter Tail Power Co. v. United States*, 410 U.S. 366 (1973); *United States v. Terminal Railroad Association*, 224 U.S. at 405; and *City of Mishawaka v. American Electric Power Co.*, 465 F.Supp. 1320, 1336 (N.D.Ind.1979), *aff'd in relevant part*, 616 F.2d 976 (7th Cir.1980).

Translation: “Google prominently displays links to its own specialized [sic] search services within its web searches and does not inform users of this favourable [sic] treatment.” Users don’t realize Google is favoring its own search services and may be steered away from “potentially more relevant competing services.” “[C]ompetitors’ results that are potentially more relevant are less visible and sometimes not directly visible to users.” “The Commission is concerned that this practice unduly diverts traffic away from Google’s competitors...”

Reaction: So what? It’s Google’s bat, ball, and playing field. Google can decide who plays and on what terms. Under *Trinko* and *linkLine*, what’s the problem?

“ – the use by Google without consent of original content from third party websites in its own vertical Web search services...”

Translation: In displaying search results describing other web sites, Google uses content from those sites without obtaining permission, such as customer reviews. When content owners object, Google says if they don’t like it, they can opt of being displayed in Google searches.

Reaction: This is certainly a concern of copyright and intellectual property law (i.e., is Google engaging in fair use), but is it a concern of antitrust law? In some sense, this might be seen as a form of reciprocal dealing, which at one time was condemned as a per se violation of Section 1, and, presumably, *a fortiori* a violation of Section 2, when engaged in by a monopolist. *Blackburn v. Sweeney*, 53 F.3d 825, 828 (7th Cir., 1995); *Key Enterprises of Delaware, Inc. v. Venice Hosp.*, 919 F.2d 1550, 1561-62 (11th Cir. 1990); *Battle v. Lubrizol Corp.*, 673 F.2d 984, 987-88 (8th Cir. 1982). Today, one can search in vain for even a rule of reason violation of Section 1 premised on reciprocal dealing. In essence, Google says that if a web site wants to appear in a Google search without charge, the web site must give Google a royalty-free copyright license to its content. Query whether this amounts to the wilful exercise of monopoly power.

“ – agreements that de jure or de facto oblige websites owned by third parties...to obtain all or almost all of their online search advertisement requirements from Google...”

Translation: Google is accused of requiring online publishers, such as newspaper web sites, to agree that the only ads on their sites for Internet search services will be Google’s, and that they will not accept ads from search services competing with Google.

Reaction: At last, an alleged violation that clearly falls within traditional United States antitrust jurisprudence. This is classic Microsoft naughty behavior, invoking shades of *Lorain Journal Co. v. United States*, 342 U.S. 143 (1951).

“ – contractual restrictions on the management and transferability of online search advertising campaigns across search advertising platforms...”

Translation: When companies sign up with Google to have ads linking to their web sites appear next to Google search results, Google allegedly requires them to do this exclusively with Google and makes the transferability of such advertising to other search engines difficult.

Reaction: Again, a traditional *Lorain Journal* type of misconduct easily falling within Section 2.

C. Google's Proposed Commitments

In response to the Commission's concerns, Google offered a number of "commitments," while denying that it had engaged in any of the subject practices. *See*, "Commission seeks feedback on commitments offered by Google to address competition concerns." Available at: <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52013XC0426%2802%29>.

First, Google basically rolled over on the last two alleged violations and abjured exclusive dealing practices going forward. Google committed that it would "no longer include in its agreements with publishers any provisions or impose any unwritten obligations that would, de jure or de facto, require publishers to source their requirements for online search advertisements exclusively from Google." It also agreed that it would "cease to impose any written or unwritten obligations...that will prevent advertisers from porting and managing search advertising campaigns across Google's AdWords and non-Google advertising services." One would expect such commitments from a business with the slogan "Do no evil," and the belief "You can make money without doing evil," which appears as number six in Google's list of "Ten things we know to be true." (<http://www.google.com/intl/en/about/company/philosophy/>)

What is more interesting and problematical is Google's response to the first two alleged violations.

As regards its alleged discriminatory treatment in favor "of links to Google's own vertical Web search services as compared to competing vertical Web search services," Google committed to a program of full disclosure.

Google will label links to Google's own vertical Web search services that are subject to a favorable placement in Google's horizontal Web search results. The label shall inform users that the links to Google's own vertical Web search services have been added by Google to provide access to its vertical Web search services, so that users do not confuse links to Google's own vertical Web search services with links to other horizontal Web search results. Where applicable, the label shall also inform users of where, in Google's horizontal Web search results, they can find links to alternative vertical Web search services.

Also, "Where applicable, Google will also distinguish links to Google's own vertical Web search services from other horizontal Web search results, so that users are made aware of their different nature." Finally, Google will display links to three web search competitors in search results where it displays its own vertical search services under certain circumstances.

Reaction: This appears to be application of the essential facilities doctrine in spades, something no United States regulator would request and no United States Court would require.

As regards the use of original content from web sites referenced in search results, "Google will offer third party websites a Web-based opt-out from the use of all content

crawled from their site in Google’s vertical Web search services,” without prejudice to their ranking in Google search results. Google will further allow competing search services “the possibility to mark certain categories of information in such a way that such information will not be indexed or used by Google.”

Also, “Google will maintain for newspaper publishers...mechanisms to enable them to control, on a web page by web page basis, the display of their content in Google News.”

Reaction: What does this have to do with regulation of competition, or even false advertising? Unless one is resurrecting the old reciprocal dealing cases, this commitment is a head-scratcher in terms of how Google is avoiding a restriction a competition authority or court could impose on its business.

Following a recitation of Google’s above-described commitments, the Commission invited comments.

D. The Commission’s Questions and Answers Sheet

Concomitant with its release of Google’s proposed commitments on April 25, 2013, the Commission also released its own Questions and Answers, which illuminate its thinking on the challenged practices. See, “Commission seeks feedback on commitments offered by Google to address competition concerns – questions and answers.” Available at: http://europa.eu/rapid/press-release_MEMO-13-383_en.htm. Some of the more interesting sections include:

Why does the Commission decide to intervene in such a fast-moving market, where the pace of innovation is rapid and a company that may be dominant today could be challenged or even replaced by another tomorrow?

In high-tech markets in particular, network effects may lead to entrenched market positions. Google has had a strong position in web search in most European countries for a number of years now. It does not seem likely that another web search service will replace it as European users’ web search service of choice.

Reaction: This certainly poses a key question, but whether the answer is convincing is another matter. The answer assumes the conclusion that Google’s “strong position in web search in most European countries for a number of years” is the result of network effects or has produced network effects that will ensure its continuance. Both propositions are far from self-evident and would benefit from further explication, which the Commission unfortunately fails to provide.

The US Federal Trade Commission investigated the way Google displays links to its specialised search services in its web search results and concluded that there was no competition issue with it. Why does the Commission come to a different view?

The factual and legal environments are different in the US and Europe. In particular, Bing and Yahoo represent a substantial alternative to Google in web search in the USA: their combined market share is around

30%. In contrast, Google has been holding market shares well above 90% in most European countries for a number of years. Web sites therefore rely more on traffic from Google in Europe than in the USA. Given the resulting commercial significance of Google for specialized search services, the way Google presents its web search results therefore has a much more significant impact on users and on the competitive process in Europe than it does in the USA.

Reaction: This is another very important question, with, regrettably, a rather facile and far from persuasive answer, which purports to reconcile a significant divergence in enforcement. The Commission's point seems to be that it is consistent enforcement policy for the Commission to be concerned with a 90 percent market share and for U.S. regulators to be unconcerned with a 70 percent market share. Who is kidding whom? Global antitrust enforcement can only benefit from a candid response here that the Commission believes the U.S. regulators to be wrong, and why. It is not credible to posit or imply an enforcement continuum on which certain behavior is antitrust-compliant with a 70 percent market share and violative with a 90 percent share.

Is the Commission not seeking to protect competitors rather than consumers?

The Commission does not act to protect competitors as such, but to preserve the competitive process for the benefit of consumers. It acts only when there is harm to competition with negative effects on consumers, in particular in terms of reduced choice and less innovation.

In particular, the Commission is concerned that the way in which Google currently presents its web search results limits the ability of European users to find their way to specialised search services competing with Google which contain information relevant to their query. Many such services might be potentially very innovative and Google's practices could therefore be limiting European consumers' opportunities to benefit from such innovative services. At the same time, it is for users to decide whether they wish to visit these sites based on their merits.

Reaction: This is the most disappointing of all the Commission's responses, in which it punts on one of the most important of all global antitrust enforcement questions. The question itself goes to the very heart of the underlying purposes of antitrust laws, wherever in the world they exist. It is impossible "to preserve the competitive process for the benefit of consumers" without protecting competitors. Once the focus of the discussion shifts to protecting "consumers," the objective becomes lower prices. When lower prices are the desideratum of antitrust law, alleged efficiencies attain importance. When alleged efficiencies are an antitrust objective, concentrations of economic power are not only tolerated, but welcomed and approved. When concentrations of economic power are sanctioned, concentrations of political power follow. When those concentrations involve communications and information-sharing media, critical instrumentalities of functioning democracy are foreclosed and stifled. If any part of the world is familiar with these principles, it is Europe, which has learned its lessons the hard way, if not the hardest way. The Commission, as the competition authority and spokesperson for Europe, ought to say so, and not pay lip service to the myth of the market and consumer welfare that all too often informs United States antitrust jurisprudence.

In fairness, the Commission invokes consumer choice in its discussion, which is a surrogate for protecting competitors without using those terms. Nonetheless, addressing the issue directly is greatly preferable to circumlocution. Indeed, as will be discussed, what is really behind the Google proceedings is exactly this view of competition regulation. It ought to be addressed for what it is, because it is ultimately a defensible view of the purposes and proper application of antitrust law.

What is the problem with Google using snippets of third party sites? If Google is infringing IP rights, can't third parties sue Google?

Intellectual Property law and competition law are two different bodies of law. Compliance with one does not necessarily imply compliance with the other, just like breaching one does not necessarily imply breaching the other.

The Commission has analysed Google's practice from the point of view of competition law. If Google's market position in web search gives it the ability to copy and use all relevant information available on the web on its own specialised search services, users may no longer have incentives to visit competing services. Competitors of Google may lose the incentive to innovate or invest in the generation of original content. This competition concern arises whether or not the information copied and used by Google is covered by IP rights.

Reaction: This is another rather oblique, if not obscure explanation by the Commission. It should say instead that it believes that "Google's market position in web search gives it the ability to copy and use all relevant information available on the web" without permission, while competing search engines lack the market power to use content from web sites without appropriate licenses. The Commission should also make express reference to the law of reciprocal dealing, inasmuch as this is essentially what is supposedly happening: Google is requiring in substance a royalty-free copyright license from any web site that wishes to appear in Google search results. Eliminating such opacity in the Commission's explanations, though inviting criticism, would at least intelligibly frame the issues for the open and honest dialogue they deserve.

Is Google benefitting from special treatment by the Commission?

The Commission is exploring the possibility of a settled outcome with Google on its four competition concerns. The possibility for a company subject to an antitrust investigation to propose commitments which the Commission can decide to make legally binding was established in 2004 by Article 9 of the EU Antitrust Regulation (Regulation 1/2003). Since this possibility was established, the Commission has taken 30 decisions making such commitments legally binding on companies.

Using this possibility may be particularly useful to swiftly restore competitive conditions on a market, for example in fast-moving markets in the IT sector. In particular, the Commission has accepted commitments by Microsoft (see IP/09/1941), Apple (see IP/12/1367) and IBM (see IP/11/1539) and turned them into legally binding obligations.

Reaction: This time the Commission fairly poses and fairly answers the question. Google is entitled to the same procedures as other targets of EC investigations, and the Commission is rightfully providing them.

What are the next steps?

The commitments are now subject to a market test of one month. Complainants, third parties and members of the public are therefore able to comment on the commitments, and the extent to which they address the Commission's four concerns.

If following the market test, the commitments form the basis for a satisfactory solution to the Commission's competition concerns, the Commission may make them legally binding on Google by way of a Commitments Decision (so-called "Article 9 procedure"). Such a decision does not conclude that there is an infringement of EU antitrust rules, but would legally bind Google to respect the commitments offered. If a company breaks such commitments, the Commission can impose a fine of up to 10% of its annual worldwide turnover.

The Commission will study all feedback very carefully and will take it into account in its analysis of whether Google's proposals address the four competition concerns. The Commission will in particular assess whether the commitments may need to be improved to adequately address the four competition concerns that have been identified.

This is in fact the procedure that ensued.

E. The Commission Accepts Google's "Improved Commitments Proposal"

On February 5, 2014, the Commission announced that it had accepted "an improved commitments proposal" from Google with respect to "the favourable [sic] treatment, within Google's horizontal Web search results, of links to Google's own vertical Web search services as compared to competing vertical Web search services," which was the first of the Commission's four competitive concerns. Available at: http://europa.eu/rapid/press-release_IP-14-116_en.htm; http://europa.eu/rapid/press-release_SPEECH-14-93_en.htm; and http://europa.eu/rapid/press-release_MEMO-14-87_en.htm?locale=en.

According to the Commission, Google agreed "to guarantee that whenever it promotes its own specialised search services on its web page (e.g., for products, hotels, restaurants, etc.), the services of three rivals, selected through an objective method, will also be displayed in a way that is clearly visible to users and comparable to the way in which Google displays its own services." This "improved" commitment would apply "not only for existing specialised search services, but also to changes in the presentation of those services and for future services."

The Commission's acceptance of the Google proposal means that the Google investigation will end, unless discussions with complainants convince the Commission that it needs to reopen the matter. The procedure, as explained by Vice President in charge of competition policy, Joaquín Almunia, will be as follows:

...we will now engage with all the 18 formal complainants in this case by outlining transparently and in detail in pre-rejection letters the reasons why we believe Google's final offer can now address the competition concerns that have been identified. Those letters will also explain why we do not believe that other issues raised by complainants are founded.

I will analyse thoroughly the feedback they will provide and only after that will I propose to the College of Commissioners to adopt a final decision. This process will take a number of months.

The Commission also provided a number of screen shots to illustrate the changes to Google's search displays that will result from Google's commitments:

1) Shopping

The Google page today:

Google 

Web Shopping Maps Images Videos More Search tools

About 8,240,000 results (0.20 seconds)

Shop for gas grills on Google Sponsored

				
GP-Grill Gas Grill - black/... £141.96 www.Ambient...	Falcon Dominator ... £888.00 Catering Appl...	Outback Omega 250 ... £129.00 Outback Direct	Blue Seal Cobra CS9 ... £897.60 Carlton Sales	Burco 444449459 ... £850.50 e-tradecounter

Ads

Gasbarbecue nodig?
www.vanhattlemhoreca.nl/ - Diverse BBQ's in de aanbieding
Bestel veilig en snel online

Catering Gas Grills
www.nisbets.co.uk/Cooking-Machines - ★★★★★ 10,305 seller reviews
Top Quality Cooking Machines
At low Prices. Free P&P Available

Buy a Gas Grill 2014 - Gas Grill Ratings - Gas Grill Reviews
bbq.about.com/cs/grills/bb/aabyb042503.htm - Before you run to the hardware store to buy a new gas grill you should know that there are a lot more grills out there than you'll find in one store. I've broken ...
Top 10 Gas Grills between ... - Top 10 Gas Grills under \$250 - Gas Grills under \$250
More by Derrick Riches - in 1,156 Google+ circles

barbecueok@gas BBQs
www.barbecueok.com/gas - Gasbarbecues met Extra Veel Smaak!
Ontdek nu de barbecuek Gas Grills.

Top Gas grill Reviews | Best Gas grill - Consumer Reports
www.consumerreports.org > Home > Home & garden - Looking for the Best gas grill? Consumer Reports has honest Ratings and Reviews on gas grills from the unbiased experts you can trust.

Gas grills
www.bestlist.nl/gasbarbecues - Gasbarbecues nu al vanaf €39.95!
Keuze uit ruim 113 gasbarbecues

Weber.com - Weber® Grills - Gas
www.weber.com/grills/category/gas - Manufacturer of gas and charcoal grills, parts and accessories. Features recipes, service and dealer locations.
Grills - Gas - Parts & Support - Charcoal

Gas Barbecues 70% off RRP
www.outdoorlivingworld.co.uk/Gas_bbq - Huge Range Of Gas Barbecues
At Great Prices - Buy Online & Save

Gas-Grills im Angebot
www.nextag.de/Gas-Grills - Erstklassige Gas-Grills in vielen Varianten: Hier zum Niedrigpreis!

Screenshots with implementation of commitments:

Google

Web Images Maps Shopping More Search tools

About 44,000,000 results (0.18 seconds)

Sponsored

Google Shopping results



GP-Grill Gas Grill - black/ ...
£141.96
www.Ambie...



Blacktop 360 Party Hub Gas
£299.00
Garden Gift...



OUTBACK 57cm ...
£127.99
Outback ...

Alternatives



Supaprize
Best stock of gas grills
from £150.00



Kelkoo
Great deals on gas grills
from £129.00



Shopzilla
Best prices on gas grills
from £180.00

Ads

Gas Grill at Amazon
www.amazon.com/patio
★★★★★ 336 reviews for amazon.com
Low Prices on **Gas grill**
Free Shipping on Qualified Orders.

Cheap Gas Grill
www.groupon.com/
★★★★★ 61 reviews for groupon.com
Save 50-90% Every Day
Thousands of Deals to Enjoy

gas grill
www.appliancesconnection.com/Grills
★★★★★ 3,550 seller reviews
1 (800) 299 9470
Huge Variety, Free In-Home Delivery
Price Match Promise on All **Grills**

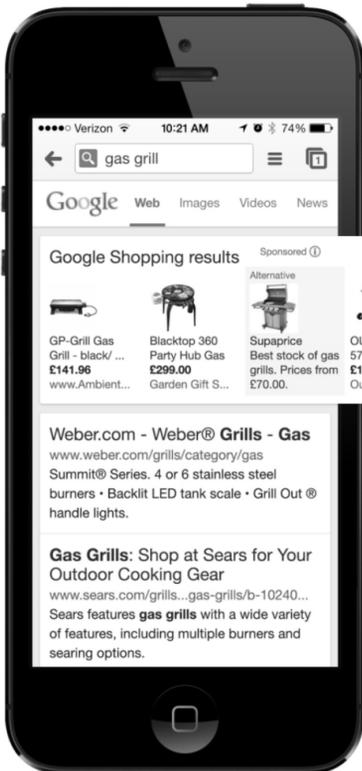
BBQ Grills
www.wholesalepatiotore.com/
1 (888) 611 7227
Quality **BBQ Grills**
Come See Our Store Today!

DCS Gas Grills
www.ajmadison.com/
★★★★★ 1,074 seller reviews

Weber.com - Weber® Grills - Gas
www.weber.com/grills/category/gas
Gas Grills. Grill Shopping Tools. 0 Grill Comparison Grill Finder. Spirit® ... 4 or 6 stainless steel burners • Backlit LED tank scale • Grill Out® handle lights.

Gas Grills: Shop at Sears for Your Outdoor Cooking Gear
www.sears.com/grills-outdoor-cooking-gas-grills/b-1024073
Sears features **gas grills** with a wide variety of features, including multiple burners and searing options. Entertain outdoors year round with **gas grills**.

Natural Gas Grills - Home Depot
www.homedepot.com/...Grills-Grill...Gas-Grills...Gas-Grills/N-5yc1vZc5t...
Natural **Gas Grills** - **Gas Grills** - Grills & Grill Accessories - Outdoors at The Home Depot.



Google Shopping results Sponsored



GP-Grill Gas Grill - black/ ...
£141.96
www.Ambie...



Blacktop 360 Party Hub Gas
£299.00
Garden Gift S...



Supaprize
Best stock of gas grills. Prices from £70.00.



OUTBACK 57cm ...
£127.99
Outback Direct



Anthony Worrall Thompson 2 ...
£250.00
Ebay



Kelkoo
Great deals on gas grills, starting at £123.



Swiss Grill Icon I430 3 Burner
£299.00
bedsbbq.co.uk



Falcon Dominator Plus G3532 ...
£888.00
Catering Appia...



Shopzilla
Best prices on gas grills from £63.99.

Weber.com - Weber® Grills - Gas
www.weber.com/grills/category/gas
Summit® Series. 4 or 6 stainless steel burners • Backlit LED tank scale • Grill Out® handle lights.

Gas Grills: Shop at Sears for Your Outdoor Cooking Gear
www.sears.com/grills...gas-grills/b-10240...
Sears features **gas grills** with a wide variety of features, including multiple burners and searing options.

2) Local search

The Google page today:

Google

Web Images Maps Vidéos Actualités Plus Outils de recherche

Environ 217 000 000 résultats (0,25 secondes)

Café de France
www.cafe-france.fr
2.7 ★★★★★ 16 avis de Google

Café de Paris
plus.google.com
3.8 ★★★★★ 25 avis de Google

Café de la Paix
www.cafedelapaix.fr
3.9 ★★★★★ 624 avis de Google

Les Deux Magots
www.lesdeuxmagots.fr
3.8 ★★★★★ 625 avis de Google

Delaville Café
www.delavillecafe.com
3.6 ★★★★★ 91 avis de Google

Café Beaubourg
www.beaumary.com
3.6 ★★★★★ 163 avis de Google

Le café A
www.facebook.com
3.9 ★★★★★ 61 avis de Google

- A 12 Place d'Italie
Paris, France
+33 1 43 31 19 86
- B 10 Rue de Buci
Paris, France
+33 1 46 34 84 11
- C 5 Place de l'Opéra
Paris, France
+33 1 40 07 36 36
- D 6 Place Saint-Germain
des Près
Paris, France
+33 1 45 48 55 25
- E 34 Boulevard de Bonne
Nouvelle
Paris, France
+33 1 48 24 48 09
- F 100 Rue Saint-Martin
Paris, France
+33 1 48 87 63 96
- G 148 Rue du Faubourg
Saint-Martin
Paris, France
+33 9 81 29 83 38

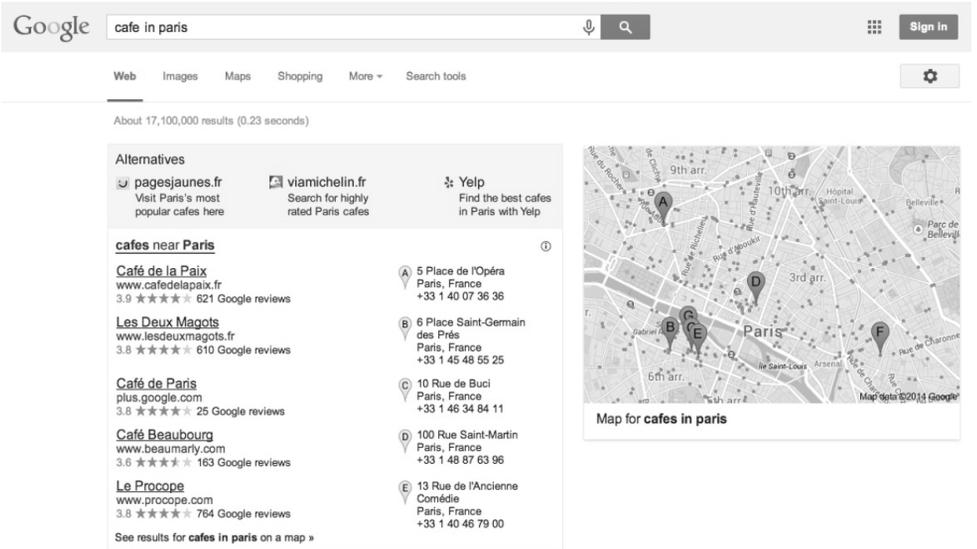
Afficher les résultats pour "cafe in paris" sur la carte »

10 of the best cafes in Paris | Travel | The Guardian
www.theguardian.com › Travel › Paris city guide › Traduire cette page
5 mai 2011 - Cult food blogger Clotilde Dusoulier selects her pick of Paris's famous cafe scene.

Paris restaurants and cafés – Time Out Paris
www.timeout.com/paris/en/restaurants-cafes
Paris restaurant guide, including the latest restaurant reviews and features. Discover the best restaurants and cafes for drinking and dining in Paris with Time Out ...

Plan de "cafe in paris"

Screenshots with implementation of commitments:

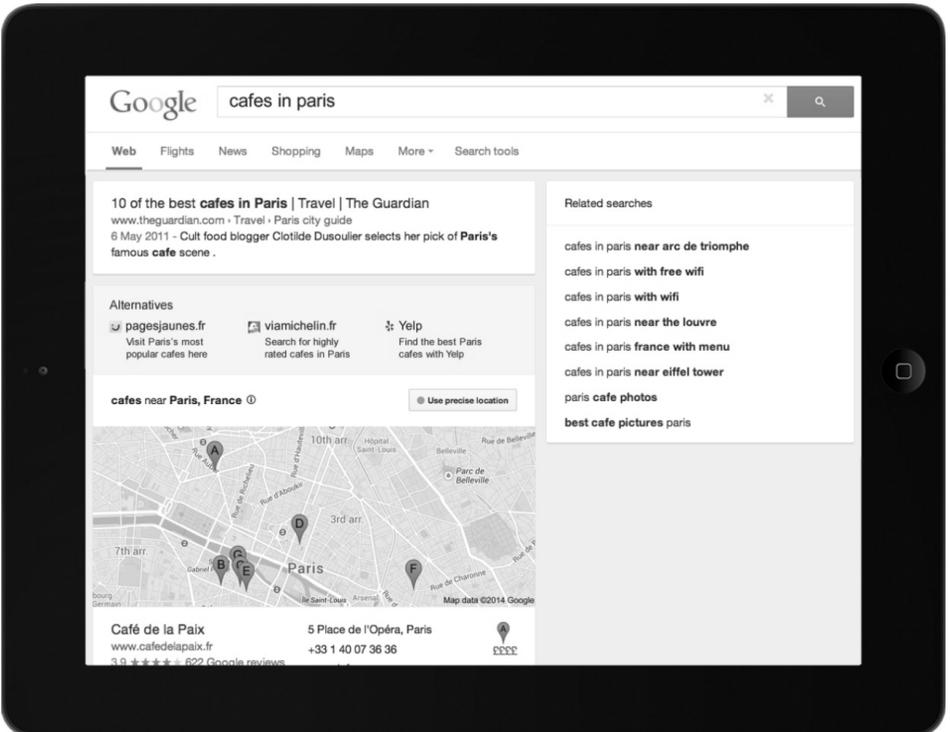


10 of the best cafes in Paris | Travel | The Guardian

www.theguardian.com › Travel › Paris city guide
6 May 2011 - Cult food blogger Clotilde Dusoulier selects her pick of Paris's famous cafe scene.

Paris restaurants and cafés – Time Out Paris

www.timeout.com/paris/en/restaurants-cafes
Paris restaurant guide, including the latest restaurant reviews and features. Discover the best restaurants and cafes for drinking and dining in Paris with Time Out ...
The 50 best restaurants in Paris - Paris's best cheap eats - Latest restaurant reviews



Defending its decision, the Commission set forth a statement from Mr. Almunia:

My mission is to protect competition to the benefit of consumers, not competitors. I believe that the new proposal obtained from Google after long and difficult talks can now address the Commission's concerns. Without preventing Google from improving its own services, it provides users with real choice between competing services presented in a comparable way; it is then up to them to choose the best alternative. This way, both Google and its rivals will be able and encouraged to innovate and improve their offerings. Turning this proposal into a legally binding obligation for Google would ensure that competitive conditions are both restored quickly and maintained over the next years.

Reaction: The first sentence from Mr. Almunia is regrettable, a sop to discredited Chicago school antitrust dogma and the myth of the market. It is really time for someone in authority to say that competition cannot be protected without protecting competitors. United States antitrust jurisprudence is on the unfortunate path to protecting competition all the way to monopoly, where there are no competitors, or oligopoly, where competitors exist in name only, and not only economic power, but political power resides in one or a few firms. This is obviously not what Congress or the European Union intended when each enacted its antitrust laws.

The interesting question is whether the Commission obtained a sufficient commitment from Google to remedy the perceived problem of anticompetitive discrimination by Google in its search displays, or whether the Commission accepted the best concession it could get from Google because the Commission perceived it was on shaky legal ground in charging an abuse of dominance. The Commission obviously takes the former view, but the 18 formal complainants in the proceedings may not concur.

The likely upshot, however, is that the Google investigation is all but over, with the Commission proposing to the College of Commissioners a final order adopting the Google commitments, and the College of Commissioners accepting it, notwithstanding howls of protests from Google's competitors.

III. CONCLUSION

At the end of the day, half of the Google settlement is unremarkable from an antitrust standpoint, while the other half is quite remarkable, not necessarily for the results achieved, but for the purpose and intent of the Commission in pursuing difficult and important issues all too infrequently linked to present-day antitrust enforcement.

First the unremarkable part. Google should never have expected to be able to require publishers to source their requirements for online search advertisements exclusively from Google, or to prevent advertisers from porting and managing search advertising campaigns across Google's AdWords and non-Google advertising services. These types of exclusive dealing arrangements by firms with monopoly or market power have long been forbidden under established antitrust precedents. Google's voluntary commitment to abjure such practices is further proof, if any is needed, that discretion is the better part of valor. Litigating the defensibility of these practices would have been a losing battle.

Now the remarkable part. In challenging Google's use of third party web site content and favoring its own search engines in search displays, the Commission entered territory long abandoned by United States regulators and Courts: reciprocal dealing and forced sharing via essential facilities doctrine. Just as remarkably, the Commission secured commitments from Google addressing both issues, though perhaps not everything the 18 formal complainants wanted.

In doing so, the Commission struck a blow for divergence, not convergence – for which the Commission deserves applause and commendation. The wielding of great economic power, whether singly or through consolidation, is an appropriate subject for antitrust regulators; and, as was once the conventional wisdom in U.S. antitrust jurisprudence, those possessing such power are forbidden in many respects from undertaking conduct open to those without it. When such power exists in the sphere of communications media or channels of information dissemination, greatly enhanced antitrust scrutiny is both appropriate and necessary, inasmuch as economic power in these industries inevitably leads to and goes hand in hand with political power.

Where foreclosure of economic freedom portends foreclosure of First Amendment and other political freedoms, antitrust law has, indeed has always had, an important part to play. As the Supreme Court said in *Associated Press v. United States*, 326 U.S. 1, 20 (1945),

It would be strange indeed, however, if the grave concern for freedom of the press which prompted adoption of the First Amendment should be read as a command that the government was without power to protect that freedom. ***The First Amendment, far from providing an argument against application of the Sherman Act, here provides powerful reasons to the contrary.*** That Amendment rests on the assumption that the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public, that a free press is a condition of a free society. Surely a command that the government itself shall not impede the free flow of ideas does not afford non-governmental combinations a refuge if they impose restraints upon that constitutionally guaranteed freedom. ***Freedom to publish means freedom for all and not for some. Freedom to publish is guaranteed by the Constitution, but freedom to combine to keep others from publishing is not.*** Freedom of the press from governmental interference under the First Amendment does not sanction repression of that freedom by private interests. ***The First Amendment affords not the slightest support for the contention that a combination to restrain trade in news and views has any constitutional immunity.*** [Emphasis added.]

Miami Herald Publishing Co. v. Tornillo, 418 US 241, 251-52 (1974); *Citizen Publishing Co. v. United States*, 394 US 131, 139-40 (1969); *Kansas City Star Company v. United States*, 240 F. 2d 643, 665-66 (8th Cir. 1957):

The theory of equal justice under law does not admit to the proposition that there is one brand of justice for some people and a different brand for others. Publishers of newspapers must answer for their actions in the same manner as anyone else. A monopolistic press could attain in tremendous measure the evils sought to be prevented by the Sherman Anti-Trust Act. Freedom to print does not mean freedom to destroy. To

use the freedom of the press guaranteed by the First Amendment to destroy competition would defeat its own ends, for freedom to print news and express opinions as one chooses is not tantamount to having freedom to monopolize. To monopolize freedom is to destroy it.

Hale v. FCC, 425 F. 2d 556, 561 (D.C. Cir., 1970) (“It is also becoming increasingly obvious that application of antitrust doctrines in regulating the mass media is not solely a question of sound economic policy; it is also an important means of achieving the goals posited by the first amendment.”); *Herbert v. Lando*, 568 F. 2d 974, 977 (2nd Cir., 1977) (“Even the one governmental control – antitrust legislation – that has long been applied to the press and does not contravene the First Amendment has been justified by its instrumental role in insuring the broad distribution of news.”).

The strictures of the foregoing decisions have unfortunately been forgotten or ignored of late in the United States, as regulators have regrettably blessed Comcast’s acquisition of NBC Universal and may be about to do the same for its grab of Time Warner cable. All credit therefore goes to the European Commission for its Google investigation. Even if the Commission did not articulate the rationale for its actions in language specifically redolent of the foregoing decisions, the intent and spirit informing the Google proceedings are undeniable, important, and essential to the maintenance of both economic and political liberty.

TECHNOLOGY-ASSISTED REVIEW: THE JUDICIAL PIONEERS

Paul E. Burns and Mindy M. Morton**
Procopio, Cory, Hargreaves & Savitch LLP
San Diego, CA
Menlo Park, CA
Scottsdale, AZ*

I. INTRODUCTION

It has been a little over two years since the first, ground-breaking court opinion addressing technology-assisted review was issued by United States Magistrate Judge Andrew J. Peck of the Southern District of New York.¹ Indeed, Judge Peck referred to this game-changing technological development in e-discovery by the more accurate phrase “computer-assisted review.” Other synonymous phrases used during the past two years include “machine assisted review” and “artificial intelligence assisted review.” More recently, the term “predictive coding” has eclipsed all of the other phrases in the case law, but as explained below, predictive coding appears to be viewed as a form of technology-assisted review, rather than as a synonym.

Two of the most recognized experts in the field, Attorney Maura R. Grossman and Professor Gordon V. Cormack, published the seminal law review article on this subject in 2011, calling it “Technology-Assisted Review.”² Moreover, Ms. Grossman and Professor Cormack have followed their seminal work with the recently published *Grossman-Cormack Glossary of Technology-Assisted Review* (“the TAR Glossary”).³ Given that we view the TAR Glossary as the most authoritative definitional source on the subject, we adopt the phrase “Technology-Assisted Review” (abbreviated as “TAR”) throughout this paper, with the caveat that we also heed Shakespeare’s advice regarding the utility of names.⁴

* B.S. Comp. Sci., *magna cum laude*; J.D., *magna cum laude*, Boston College Law School. Adjunct Professor of Law, Pepperdine University School of Straus Institute for Dispute Resolution; Adjunct Professor of Law, Arizona State University Sandra Day O’Connor College of Law; Director and Executive Committee Member, The Sedona Conference; Attorney, Procopio, Cory, Hargreaves & Savitch, LLP, Menlo Park, CA, San Diego, CA and Scottsdale/Phoenix, AZ.

** B.A., *highest distinction*, University of California, Berkeley; J.D., Yale Law School; Attorney, Procopio, Cory, Hargreaves & Savitch, LLP, Menlo Park, CA.

1 *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012).

2 Maura R. Grossman and Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 12 RICH J.L. TECH. 1 (2011) (see also Foreword by John M. Facciola, U.S. Magistrate Judge). This article has been cited in at least eight reported judicial decisions to date.

3 Maura R. Grossman and Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review*, 7 FED. COURTS L. REV. 1 (2013) (hereinafter cited as “TAR Glossary”).

4 See William Shakespeare, *Romeo and Juliet*, Act II, Scene 2 (“What’s in a name? that which we call a rose By any other name would smell as sweet.”).

Since February, 2012, when Judge Peck boldly went where no judge had gone before in writing an opinion addressing TAR, we found several pioneering judges who have followed him into the final frontier. Indeed, we now have what could be considered to be a critical mass of published opinions that can collectively provide some badly needed TAR guidance to the bench and bar. Whether counsel is meeting and conferring regarding the potential use of TAR, or a court is considering the efficacy of a proposed TAR protocol, these pioneering judicial opinions are a must read.

To ease the load on lawyer and judge alike, we have endeavored to create this compilation and summary of some of the currently available published opinions addressing TAR. We salute the pioneering authors of these decisions, and while they may not all agree in approach or outcome, each and every one of them have contributed to moving the law of TAR forward in a reasoned and just way.

We begin with the definition of TAR and its younger sibling, predictive coding, as well as Judge Peck's seminal opinion in *Da Silva Moore*. We will then follow with the writings of Judge Peck's fellow judicial pioneers.

II. TECHNOLOGY-ASSISTED REVIEW DEFINED

The TAR Glossary defines TAR as follows:

Technology-Assisted Review (TAR): A process for Prioritizing or Coding a Collection of Documents using a computerized system that harnesses human judgments of one or more Subject Matter Expert(s) on a smaller set of Documents and then extrapolates those judgments to the remaining Document Collection. Some TAR methods use Machine Learning Algorithms to distinguish Relevant from Non-Relevant Documents, based on Training Examples Coded as Relevant or Non-Relevant by the Subject Matter Experts(s), while other TAR methods derive systematic Rules that emulate the expert(s)' decision-making process. TAR processes generally incorporate Statistical Models and/or Sampling techniques to guide the process and to measure overall system effectiveness.⁵

While some experts and E-Discovery vendors may place certain emphasis on portions of the above definition or may add other aspects, the Grossman-Cormack definition captures the essence of the current thinking on what Technology-Assisted Review comprises. Indeed, the most significant aspect of the TAR Glossary's definition is that it encompasses multiple TAR methods and processes.

Grossman and Cormack explain predictive coding as follows:

“Predictive Coding: An industry-specific term generally used to describe a Technology-Assisted Review process involving the use of a Machine Learning Algorithm to distinguish Relevant from Non-Relevant Documents, based on Subject Matter Expert(s)' Coding of a Training Set of Documents. See Supervised Learning and Active Learning.”⁶

⁵ *TAR Glossary*, 7 FED. COURTS L. REV. at 32.

Thus, the TAR Glossary establishes that predictive coding is not precisely synonymous with TAR, but rather describes an industry-specific term generally used to describe a particular TAR process. Indeed, predictive coding is a species of the genus TAR. Moreover, depending upon the particular e-discovery vendor's flavor of predictive coding, the term could be used to identify a particular process that differs from other processes applying the same term. In other words, there may be multiple sub-species of predictive coding.

III. JUDGE PECK'S EXPEDITION INTO THE FINAL FRONTIER: *DA SILVA MOORE V. PUBLICIS GROUPE*

The first published judicial decision addressing TAR was written by Judge Peck in the case of *Da Silva Moore v. Publicis Groupe*.⁷ *Da Silva Moore* was a sex discrimination case in which the defendant, "one of the world's 'big four' advertising conglomerates," sought to utilize TAR to reduce the massive amount of electronically stored information ("ESI") in its possession that would otherwise require human review. Judge Peck, quoting his seminal article *Search, Forward: Will manual document review and keyword searches be replaced by computer-assisted coding?*,⁸ and opting to use the term "computer-assisted review" in lieu of TAR, wrote as follows:

"My *Search, Forward* article explained my understanding computer-assisted review, as follows":

By computer-assisted coding, I mean tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (i.e., training by) a human reviewer.

Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or [small] team) who review and code a "seed set" of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer's coding. (Or, the computer codes some documents and asks the senior reviewer for feedback.)

When the system's predictions and the reviewer's coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents. Typically, the senior lawyer (or team) needs to review only a few thousand documents to train the computer.

Some systems produce a simple yes/no as to relevance, while others give a relevance score (say, on a 0 to 100 basis) that counsel can use to prioritize review. For example, a score above 50 may produce 97% of the relevant documents, but constitutes only 20% of the entire document set.

⁶ *TAR Glossary*, 7 FED. COURTS L. REV. at 26.

⁷ *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012).

⁸ Andrew J. Peck, *Search, Forward: Will manual document review and keyword searches be replaced by computer-assisted coding?*, L. TECH. NEWS (Oct. 2011), p. 25 and 29.

Counsel may decide, after sampling and quality control tests, that documents with a score of below 15 are so highly likely to be irrelevant that no further human review is necessary. Counsel can also decide the cost-benefit of manual review of the documents with scores of 15-50.

“My article further explained my belief that Daubert would not apply to the results of using predictive coding, but that in any challenge to its use, this Judge would be interested in both the process used and the results”:

[I]f the use of predictive coding is challenged in a case before me, I will want to know what was done and why that produced defensible results. I may be less interested in the science behind the “black box” of the vendor’s software than in whether it produced responsive documents with reasonably high recall and high precision.

That may mean allowing the requesting party to see the documents that were used to train the computer-assisted coding system. (Counsel would not be required to explain why they coded documents as responsive or non-responsive, just what the coding was.) Proof of a valid “process,” including quality control testing, also will be important.

. . . .

Of course, the best approach to the use of computer-assisted coding is to follow the Sedona Cooperation Proclamation model. Advise opposing counsel that you plan to use computer-assisted coding and seek agreement; if you cannot, consider whether to abandon predictive coding for that case or go to the court for advance approval.⁹

In *Da Silva Moore*, plaintiffs’ counsel did not reject defendant’s proposal to use predictive coding “to cull down the approximately three million electronic documents from the agreed-upon custodians . . .,” but rather had “multiple concerns” and wanted “clarification” regarding the process.¹⁰ Initially, Judge Peck offered the parties the following advice:

Key words, certainly unless they are well done and tested, are not overly useful. Key words along with predictive coding and other methodology, can be very instructive. I’m also saying to the defendants . . . if you do predictive coding, you are going to have to give your seed set, including the seed documents marked as non-responsive to the plaintiff’s counsel so they can say, well, of course you are not getting any [relevant] documents, you’re not appropriately training the computer.¹¹

The first dispute concerned defendant MSL’s preliminary proposal to review only the top 40,000 most relevant documents identified by the predictive coding software, at a projected cost of \$5.00 per document for a total cost of \$200,000.00. The court rejected this

⁹ *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 183-84 (S.D.N.Y. 2012).

¹⁰ *Da Silva Moore*, 287 F.R.D. at 184.

¹¹ *Da Silva Moore*, 287 F.R.D. at 185 (hearing transcript citations omitted).

proposal, pointing out that “‘where [the] line will be drawn [as to review and production] is going to depend on what the statistics show for the results,’ since ‘[p]roportionality requires consideration of results as well as costs. And if stopping at 40,000 is going to leave a tremendous number of likely highly responsive documents unproduced, [MSL’s] proposed cutoff doesn’t work.’”¹² Thus, it appears that the court would not countenance a manual review proposal that was based solely on an arbitrary number of documents and cost cutoff.

The second dispute concerned the identification of custodians’ emails to be searched. The court employed a phased custodian approach, adopting defendant’s proposal to search thirty custodians in the first phase, comprising defendant’s president, other members of the executive team, and certain managing directors. Plaintiffs wanted additional custodians whom plaintiffs claimed would be more likely to have information bearing on discriminatory practices as to preferential job duties and assignments. Exhibiting exceptional case management skills, the court ventured deep in the weeds to understand the quality of the ESI at issue and observed that these additional custodians’ “emails would be so different from that of the other custodians,” that they should not be included in the emails subjected to predictive coding review.¹³ Indeed, the court concluded that the information sought would be more efficiently extracted through deposition testimony.¹⁴ The court also approved a two-phased approach for ESI sources, leaving what appeared to be less important or uncertain sources of ESI for the second phase. The court found the authority for multi-phasing in the proportionality rule set forth in Fed. R. Civ. P. 26(b)(2)(C).¹⁵

The court also gave attention to an issue often overlooked by American courts – foreign privacy laws blocking search and retrieval of emails. Citing to *The Sedona Conference International Principles on Discovery, Disclosure and Data Protection* (2011), the Court concluded that, because the defendant’s France-based CEO’s emails would likely be covered by French privacy and blocking laws, the CEO should not be included as a first phase custodian.¹⁶

Probably the most instructive portion of the *Da Silva Moore* decision concerns the discussion of the predictive coding protocol proposed by defendant MSL and approved by the court in a February 8, 2012 hearing:

The parties agreed to use a 95% confidence level (plus or minus two percent) to create a random sample of the entire email collection; that sample of 2,399 documents will be reviewed to determine relevant (and not relevant) documents for a “seed set” to use to train the predictive coding software. ...

To further create the seed set to train the predictive coding software, MSL coded certain documents through ‘judgmental sampling.’ The remainder of the seed set was created by MSL reviewing ‘keyword’ searches with Boolean connectors (such as ‘training and *Da Silva Moore*,’ or ‘promotion and *Da Silva Moore*’) and coding the top fifty hits from those searches. MSL agreed to provide all those documents (except

12 *Id.*

13 *Id.*

14 *Id.*

15 *Id.*

16 *Da Silva Moore*, 287 F.R.D. at 186.

privileged ones) to plaintiffs for plaintiffs to review MSL's relevance coding. In addition, plaintiffs provided MSL with certain other keywords, and MSL used the same process with plaintiffs' keywords as with the MSL keywords, reviewing and coding an additional 4,000 documents. All of this review to create the seed set was done by senior attorneys (not paralegals, staff attorneys or junior associates). MSL reconfirmed that "[a]ll of the documents that are reviewed as a function of the seed set, whether [they] are ultimately coded relevant or irrelevant, aside from privilege, will be turned over to' plaintiffs. The next area of discussion was the iterative rounds to stabilize the training of the software. MSL's vendor's predictive coding software ranks documents on a score of 100 to zero, i.e., from most likely relevant to least likely relevant. MSL proposed using seven iterative rounds; in each round they would review at least 500 documents from different concept clusters to see if the computer is returning new relevant documents. After the seventh round, to determine if the computer is well trained and stable, MSL would review a random sample (of 2,399 documents) from the discards (i.e., documents coded as non-relevant) to make sure the documents determined by the software to not be relevant do not, in fact, contain highly-relevant documents. For each of the seven rounds and the final quality-check random sample, MSL agreed that it would show plaintiffs all the documents it looked at including those deemed not relevant (except for privileged documents).

Plaintiffs' vendor noted that "we don't at this point agree that this is going to work. This is new technology and it has to be proven out." Plaintiffs' vendor agreed, in general, that computer-assisted review works, and works better than most alternatives. Indeed, plaintiffs' vendor noted that "it is fair to say [that] we are big proponents of it." The Court reminded the parties that computer-assisted review 'works better than most of the alternatives, if not all of the [present] alternatives. So the idea is not to make this perfect, it's not going to be perfect. The idea is to make it significantly better than the alternatives without nearly as much cost.'

The Court accepted MSL's proposal for the seven iterative reviews, but with the following caveat:

But if you get to the seventh round and [plaintiffs] are saying that the computer is still doing weird things, it's not stabilized, etc., we need to do another round or two, either you will agree to that or you will both come in with the appropriate QC information and everything else and [may be ordered to] do another round or two or five or 500 or whatever it takes to stabilize the system.¹⁷

Notwithstanding plaintiffs having purportedly agreed to the ESI protocol discussed in the February 8th hearing, they filed objections two weeks later alleging that the protocol was essentially forced upon them by the court. Plaintiffs' first objection was based on the argument that producing counsel needed to certify the production as "complete and correct" under FRCP 26(g)(1)(A). The court explained that plaintiffs' were wrong

¹⁷ *Id.* at 186-87 (Hearing transcript citations omitted).

because that certification only applied to initial disclosures under FRCP 26(a)(1). The court noted that the provision applicable to discovery responses, FRCP 26(g)(1)(B), incorporates the proportionality principle of FRCP 26(b)(2)(C).

Plaintiffs' second objection was that the court's acceptance of defendants' proposed ESI Protocol was contrary to Federal Rule of Evidence 702. Judge Peck observed that FRE 702 and *Daubert*¹⁸ are rules for admissibility of evidence at trial and "simply are not applicable to how documents are searched for and found in discovery."¹⁹ The Court further explained that "The admissibility of specific emails at trial will depend upon each email itself (for example whether it is hearsay, or a business record or party admission), not how it was found during discovery."²⁰

The most reasonable basis for objection offered by plaintiffs in *Da Silva Moore* is the same one that perhaps all but a privileged few in the legal profession would currently have: we just don't understand what this 'black box' called predictive coding is and "there is no way to be certain if MSL's method is reliable." In an effort to mitigate this concern, the court encouraged complete transparency in defendant's process – with defendant disclosing how it coded every e-mail used in the seed set (both relevant and irrelevant).²¹ The court ultimately ruled that plaintiffs' concerns were premature, but left the door open for plaintiffs to raise concerns again during or after the predictive coding process, "(which the Court will be closely supervising)."²²

IV. JUDGE PECK'S "FURTHER ANALYSIS AND LESSONS FOR THE FUTURE"

Much like Captain James T. Kirk would record his thoughts into the Captain's Log as he travelled through space at warp speed, Judge Peck gave us the benefit of his thoughts as he launched into the new frontier of TAR. Here is a summary of some of his most compelling observations:

1. TAR "is not a magic, Staples-Easy-Button, solution appropriate for all cases. The technology exists and should be used where appropriate, but it is not a case of machine replacing humans; it is the process used and the interaction of man and machine that the courts needs to examine."²³
2. Judge Peck teaches us that "recall," a measure of completeness, is the fraction of relevant documents identified during a review, while "precision," a measure of accuracy, is the fraction of identified documents that are relevant. "The goal is for the review method to result in higher recall and higher precision than another review method, at a cost proportionate to the "value" of the case."²⁴

18 *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

19 *Da Silva Moore*, 287 F.R.D. at 188-89.

20 *Id.* at 189.

21 *Id.*

22 *Id.*

23 *Id.*

24 *Id.* at 189-90, citing Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 12 RICH. J.L. & TECH., Spring 2011, at 8-9.

3. While some lawyers still think that linear (i.e., manual) review is the gold standard, “statistics clearly show that computerized searches are at least as accurate, if not more so, than manual review.”²⁵ Moreover, Grossman-Cormack’s seminal article noted that “technology-assisted reviews require, on average, human review of only 1.9% of the documents, a fifty-fold savings over manual review,” thus establishing significant cost savings with TAR over manual review.²⁶
4. While keyword searches have earned a place in the world of ESI production, there are too many cases in which “the way lawyers choose keywords is the equivalent of the child’s game of ‘Go Fish.’” Another problem with keyword searching “is that they often find large quantities of irrelevant ESI. The court gave examples of keywords and the number of “hits” in which they resulted (e.g., “Da Silva Moore”: 201,179 hits, “training”: 165,208 hits). Manual review of so many hits would be quite costly and many would be irrelevant. Moreover, data shows that “keyword searches are not very effective.”²⁷
5. In light of the foregoing, Judge Peck concluded: “Computer-assisted review appears to be better than the available alternatives, and thus should be used in appropriate cases. **While this Court recognizes that computer-assisted review is not perfect, the Federal Rules of Civil Procedure do not require perfection.**”²⁸
6. Moreover, Judge Peck emphasized the importance of cooperation and transparency in connection with ESI protocols, pointing out that the Peck Court “strongly endorses *The Sedona Conference Cooperation Proclamation.*” A critical lesson for counsel proposing the use of TAR is found in Judge Peck’s observation that “MSL’s transparency in its proposed ESI search protocol made it easier for the Court to approve the use of predictive coding.” Judge Peck realized that not all experienced ESI counsel would be willing to be as transparent as MSL, “such transparency allows the opposing counsel (and the Court) to be more comfortable with computer-assisted review, reducing fears about the so-called ‘black box’ of the technology. ***This Court highly recommends that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process.***”²⁹

During the course of Judge Peck’s earnest study of TAR, he has come up with the following guidelines for courts in addressing TAR proposals:

25 *Da Silva Moore*, 287 F.R.D. at 190, citing Herbert L. Roitblatt, Anne Kershaw & Patrick Oot, *Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review*, 61 J. AM. SOC’Y FOR INFO. SCI. & TECH. 70, 79 (2010).

26 *Da Silva Moore*, 287 F.R.D. at 190, citing Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 12 RICH. J.L. & TECH., Spring 2011, at 43. [Emphasis supplied].

27 *Da Silva Moore*, 287 F.R.D. at 190-91, citing David L. Blair & M. E. Maron, *An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System*, 28 COMM. ACM 289 (1985) (Searchers believed they met the goals, but their average recall was just 20%). Judge Peck noted that “This result has been replicated in the TREC Legal Track studies over the past few years.”

28 *Da Silva Moore*, 287 F.R.D. at 191 (emphasis added).

29 *Da Silva Moore*, 287 F.R.D. at 192 (emphasis added).

1. It is unlikely that a court will be able to approve an advance proposed stopping point for the TAR process or based on an arbitrary number of documents and/or their relevance ratings. The stopping point can only be determined after training the TAR software and verifying the quality of the results.³⁰
2. Staging discovery, i.e., multiple phases of ESI sources and key custodians, is an effective way to control discovery costs. However, the court must be willing in appropriate cases to grant discovery extensions when necessary.³¹
3. “If you are knowledgeable about and tell the other side who your key custodians are and how you propose to search for the requested documents, opposing counsel and the Court are more apt to agree to your approach (at least as phase one without prejudice).”³²
4. It is very helpful for counsel to have their e-discovery vendors present at court hearings where ESI Protocol is to be discussed. In Judge Peck’s words, “bring your geek to court.” “It also is important for the vendors and/or knowledgeable counsel to be able to explain complicated ediscovery concepts in ways that make it easily understandable to judges who may not be tech-savvy.”³³

Finally, Judge Peck’s conclusion summarizes the importance of his *Da Silva Moore* opinion as follows:

What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review. Counsel no longer have to worry about being the “first” or “guinea pig” for judicial acceptance of computer-assisted review. As with keywords or any other technological solution to ediscovery, counsel must design an appropriate process, including use of available technology, with appropriate quality control testing, to review and produce relevant ESI while adhering to Rule 1 and Rule 26(b)(2)(C) proportionality. Computer-assisted review now can be considered judicially-approved for use in appropriate cases.³⁴

V. THE JUDICIAL PIONEERS: POST-*DA SILVA MOORE* TECHNOLOGY-ASSISTED REVIEW OPINIONS

A. *Global Aerospace Inc. v. Landow Aviation, L.P.* – Judge Chamblin

In *Global Aerospace Inc. v. Landow Aviation, L.P.*,³⁵ the court allowed defendants to use predictive coding over plaintiffs’ objections, stating:

³⁰ *Id.*

³¹ *Id.*

³² *Id.* at 193.

³³ *Id.* at 193.

³⁴ *Id.*

³⁵ 2012 WL 1431215, No. CL 61040 (Va. Cir. Ct., Apr. 23, 2012).

Having heard argument with regard to the Motion [Defendants], pursuant to Virginia Rules of Supreme Court 4:1(b) and (c) and 4:15, it is hereby ordered Defendants shall be allowed to proceed with the use of predictive coding for purposes of the processing and production of electronically stored information, with processing to be completed with 60 days and production to follow as soon as practicable and in no more than 60 days. This is without prejudice to a receiving party raising with the court an issue as to completeness or the contents of the production or the ongoing use of predictive coding.³⁶

In their brief, defendants argued that they had:

[A]n estimated 250 gigabytes (GB) of reviewable ESI from its computer systems, which could easily equate to more than two million documents. At average cost and rates of review and effectiveness, linear first-pass review would take 20,000 man hours, cost two million dollars, and locate only sixty percent of the potentially relevant documents. As one alternative, keyword searching might be more cost-effective but likely would retrieve only twenty percent of the potentially relevant documents and would require Landow to incur substantial unnecessary costs for document review. Predictive coding, on the other hand, is capable of locating upwards of seventy-five percent of the potentially relevant documents and can be effectively implemented at a fraction of the cost and in a fraction of the time of linear review and keyword searching. Further, by including a statistically sound validation protocol, Landow's counsel will thoroughly discharge the "reasonable inquiry" obligations of Rule 4:1(g).³⁷

Plaintiffs claimed in opposition that "[C]omputerized tools are supplements to the ordinary review process. No computer program is an adequate substitute for having human beings review and sort the documents."³⁸ Plaintiffs further stated that "Defendants should produce all responsive emails and other electronic documents, not just the 75%, or less, that the "predictive coding" computer program might select."³⁹ Despite plaintiffs' concerns, the court granted the protective order sought by defendants and allowed the use of predictive coding.⁴⁰

B. *National Day Laborer Organizing Network v. U.S. Immigration and Customs Enforcement Agency, et al.* – Judge Scheindlin

Judge Scheindlin, who was the original judicial pioneer of e-discovery, addressed TAR in *National Day Laborer Organizing Network v. U.S. Immigration and Customs Enforcement Agency, et al.*,⁴¹ which concerned a Freedom of Information Act request from a number of federal agencies regarding the Secure Communities program. The case examined the adequacy of the searches performed by the various agencies, and suggested that the FBI

³⁶ *Global Aerospace Inc.* 2012 WL 1431215 at *1.

³⁷ *Defendants' Memorandum in Support of Motion for Protective Order Approving the Use of Predictive Coding*, 2012 WL 1419842, No. CL 61040, *1 (Va. Cir. Ct., Apr. 9, 2012).

³⁸ *Plaintiffs' Opposition to the Landow Defendants' Motion for Protective Order Regarding Electronic Documents and "Predictive Coding"*, 2012 WL 1419848, No. CL 61040, *1 (Va. Cir. Ct., Apr. 16, 2012).

³⁹ *Id.*

⁴⁰ *Global Aerospace Inc.* 2012 WL 1431215.

⁴¹ 877 F. Supp. 2d 87 (S.D.N.Y. 2012).

might have employed “sophisticated search techniques to ensure that the manual review was actually capturing the universe of responsive documents” as “[s]uch tests would have given the Court significantly more confidence regarding the adequacy of these manual reviews.”⁴² The court further suggested that keyword searching is not effective and that:

[B]eyond the use of keyword search, parties can (and frequently should) rely on latent semantic indexing, statistical probability models, and machine learning tools to find responsive documents. Through iterative learning, these methods (known as “computer assisted” or “predictive” coding) allow humans to teach computers what documents are and are not responsive to a particular FOIA or discovery request and they can significantly increase the effectiveness and efficiency of searches.⁴³

Although the court did not require the federal agencies to use such techniques in amending responses to the FOIA requests, the court stated that “If [the parties] wish to and are able to, then they may agree on predictive coding techniques and other more innovated ways to search.”⁴⁴

C. *EORHB, Inc. v. HOA Holdings LLC – Vice Chancellor Laster*

In *EORHB, Inc. v. HOA Holdings LLC*,⁴⁵ the court, *sua sponte*, ordered:

The parties shall confer regarding a case schedule. Absent a modification of this order for good cause shown, the parties shall (i) retain a single discovery vendor to be used by both sides, and (ii) conduct document review with the assistance of predictive coding. If the parties cannot agree on a single discovery vendor with expertise in predictive coding, the parties shall each submit up to two vendor candidates to the Court.⁴⁶

In a later order, the court entered a stipulated order allowing plaintiffs to “conduct document review using traditional methods” because, “based on the low volume of relevant documents expected to be produced in discovery by [Plaintiffs], the cost of using predictive coding assistance would likely be outweighed by any practical benefit of its use.”⁴⁷

D. *Gabriel Technologies Corp. v. Qualcomm Inc. – Judge Battaglia*

In *Gabriel Technologies Corp. v. Qualcomm Inc.*,⁴⁸ the court awarded fees and costs to defendants after a summary judgment, finding that the case was “exceptional” under 35 U.S.C. § 285.⁴⁹ The court included the amount defendants expended on TAR:

The third aspect of Defendants’ requested fees is \$2,829,349.10 attributable to computer assisted, algorithm-driven document review. Defendants provide the following explanation for the resulting fees: “Over the course of this litigation, Defendants collected almost

⁴² *Id.* at 103.

⁴³ *Id.* at 109.

⁴⁴ *Id.* at 111.

⁴⁵ 2012 WL 4896670, No. 7409-VCL (Del. Ch., Oct. 15, 2012).

⁴⁶ *Id.* at *1.

⁴⁷ 2013 WL 1960621, No. 7409-VCL (Del. Ch., May 6, 2013).

⁴⁸ 2013 WL 410103, No. 08cv1992 AJB (MDD) (S.D.Cal., Feb 1, 2013).

⁴⁹ *Gabriel Technologies Corp. v. Qualcomm Inc.*, 2013 WL 410103, No. 08cv1992 AJB (MDD), *3-5 (S.D.Cal., Feb 1, 2013).

12,000,000 records – mostly in the form of Electronically Stored Information (ESI)... Rather than manually reviewing the huge volume of resultant records, Defendants paid H5 to employ its proprietary technology to sort these records into responsive and non-responsive documents.” After the algorithm determined whether documents were responsive or unresponsive to discovery requests, Black Letter attorneys reviewed the responsive documents for confidentiality, privilege, and relevance issues. For this reason, the review performed by H5 and Black Letter accomplished different objectives with the H5 electronic process minimizing the overall work for Black Letter. Again, the Court finds Cooley’s decision to undertake a more efficient and less time-consuming method of document review to be reasonable under the circumstances. In this case, the nature of Plaintiffs’ claims resulted in significant discovery and document production, and Cooley seemingly reduced the overall fees and attorney hours required by performing electronic document review at the outset. Thus, the Court finds the requested amount of \$2,829,349.10 to be reasonable.⁵⁰

E. *Chevron v. Donziger* – Judge Kaplan

In *Chevron v. Donziger*,⁵¹ the court required nonparty (and defendants’ counsel in related action in Ecuador) Patton Boggs LLP to respond to a subpoena duces tecum.⁵² In considering (and denying) Patton Boggs’ claim of undue burden and request for cost shifting, the court ruled:

At the September 2012 hearing, the Court urged the parties to analyze, in their subsequent submissions with respect to burden, whether and to what extent predictive coding could “reduce the burden and effort” required to comply with the Subpoena. Apart from one footnote, PB’s submission ignored the subject entirely. The logical inference is that PB failed to address the subject because it would not have aided its argument.⁵³

The court also noted that “Predictive coding is an automated method that credible sources say has been demonstrated to result in more accurate searches at a fraction of the cost of human reviewers.”⁵⁴

F. *In Re Biomet M2a Magnum Hip Implant Prod. Liab. Litig.* – Judge Miller

In MDL litigation titled *In re Biomet M2a Magnum Hip Implant Products Liability Litigation*, the court issued two orders concerning the use of predictive coding.⁵⁵ In the first order, the court considered the plaintiffs’ motion regarding the sufficiency of defendant’s document review process.⁵⁶ Defendant used key-word searching to cull responsive documents from 19.5 million documents down to 3.5 million.⁵⁷ Defendant removed duplicate documents, and then used predictive coding to identify relevant

⁵⁰ *Id.* at *10.

⁵¹ 2013 WL 1087236, No. 11 Civ. 0691(LAK) (S.D.N.Y. March 15, 2013).

⁵² *Id.* at *3.

⁵³ *Id.* at *32.

⁵⁴ *Id.* at *255.

⁵⁵ *In re Biomet*, 2013 WL 1729682, No. 3:12–MD–2391 (N.D. Ind. Apr. 18, 2013); 2013 WL 6405156, No. 3:12–MD–2391 (N.D. Ind. Aug. 21, 2013).

⁵⁶ *In re Biomet*, 2013 WL 1729682.

⁵⁷ *Id.* at *1.

documents.⁵⁸ Plaintiffs asserted that defendant should have to start its review over because the use of key-word searching “has tainted the process.”⁵⁹ The court ruled:

The issue before me today isn't whether predictive coding is a better way of doing things than keyword searching prior to predictive coding. I must decide whether Biomet's procedure satisfies its discovery obligations and, if so, whether it must also do what the Steering Committee seeks. What Biomet has done complies fully with the requirements of Federal Rules of Civil Procedure 26(b) and 34(b)(2). I don't see anything inconsistent with the Seventh Circuit Principles Relating to the Discovery of Electronically Stored Information. Principle 1.02 requires cooperation, but I don't read it as requiring counsel from both sides to sit in adjoining seats while rummaging through millions of files that haven't been reviewed for confidentiality or privilege.⁶⁰

In denying plaintiffs' motion, the Court determined as follows:

It might well be that predictive coding, instead of a keyword search, at Stage Two of the process would unearth additional relevant documents. But it would cost Biomet a million, or millions, of dollars to test the Steering Committee's theory that predictive coding would produce a significantly greater number of relevant documents. Even in light of the needs of the hundreds of plaintiffs in this case, the very large amount in controversy, the parties' resources, the importance of the issues at stake, and the importance of this discovery in resolving the issues, I can't find that the likely benefits of the discovery proposed by the Steering Committee equals or outweighs its additional burden on, and additional expense to, Biomet.⁶¹

Four months later, the *Biomet* court again addressed an ESI dispute between the parties.⁶² The Plaintiffs Steering Committee brought a motion to compel Defendant “to produce the discoverable documents used in the training of the ‘predictive coding’ algorithm. Biomet reveals only that the discoverable documents used in the seed set already have been disclosed to the Steering Committee; Biomet won't identify the seed set beyond that.”⁶³ The court explained that, “As I understand it, a predictive coding algorithm offers up a document, and the user tells the algorithm to find more like that document or that the user doesn't want more documents like what was offered up.”⁶⁴ The court first determined that plaintiffs could not be asking for the entire seed set, as that could include privileged or irrelevant documents.⁶⁵ However, even focusing the request to deal only with discoverable documents, all of which had been produced but not identified as part of the seed set, the court could not find authority to order defendant to comply, stating, “I'm puzzled as to the authority behind the Steering Committee's request.”⁶⁶ The plaintiffs cited to *The Sedona Conference Cooperation Proclamation*⁶⁷ in support of their position. The court considered the *Cooperation Proclamation* and the Seventh Circuit project cited by Plaintiffs, but ultimately held as follows:

58 *Id.*

59 *Id.*

60 *Id.* at *2.

61 *Id.* at *3.

62 *In re Biomet*, 2013 WL 6405156, No. 3:12-MD-2391 (N.D. Ind. Aug. 21, 2013).

63 *Id.* at *1.

64 *Id.*

65 *Id.*

66 *Id.*

67 10 Sedona Conf. J. 331 (2009 Supp.).

[N]either the Sedona Conference nor the Seventh Circuit project expands a federal district court's powers, so they can't provide me with authority to compel discovery of information not made discoverable by the Federal Rules. Still, Biomet's position is troubling. Biomet suggests no way in which telling the Steering Committee which of the documents already produced were in the seed set would harm it. Based on what I have been given in the parties' memoranda, Biomet is right that it doesn't have to identify the seed set, but the Steering Committee is right that Biomet's cooperation falls below what the Sedona Conference endorses. An unexplained lack of cooperation in discovery can lead a court to question why the uncooperative party is hiding something, and such questions can affect the exercise of discretion. But I don't have any discretion in this dispute. I won't order Biomet to reveal which of the documents it has disclosed were used in the seed set, but I urge Biomet to re-think its refusal.⁶⁸

G. *Gordon v. Kaleida Health and Hinterberger v. Catholic Health System, Inc. – Judge Foschio*

In *Gordon v. Kaleida Health* and *Hinterberger v. Catholic Health System, Inc.*, Magistrate Judge Foschio of the Western District of New York addressed predictive coding in companion cases concerning New York wage and hour laws.⁶⁹ The plaintiffs and defendants were represented by the same counsel in both cases, and Judge Foschio issued two decisions in each case addressing ediscovery issues and predictive coding. *Gordon I* and *Hinterberger I* concerned a dispute over Plaintiffs' e-discovery and predictive coding consultant.⁷⁰ The dispute over the vendor significantly delayed the parties' predictive coding protocol, and *Gordon II* and *Hinterberger II* concern plaintiffs' motion to either compel defendants to meet and confer or to adopt plaintiffs' predictive coding protocol.⁷¹ The court observed:

For well-over a year, the parties have attempted, without success, to agree on how to achieve a cost-effective review of Defendants' voluminous e-mails, estimated at 200-300,000 using a key-word search methodology. At the last of a series of ESI discovery status conferences with the court, ... the court expressed dissatisfaction with the parties' lack of progress toward resolving issues related to completion of review and production of Defendants' e-mails using the key-word search method, and pointed to the availability of predictive coding, a computer assisted ESI reviewing and production method.⁷²

⁶⁸ *Id.* at *2.

⁶⁹ See *Gordon v. Kaleida Health*, 2013 WL 2250506, No. 08-CV-378S(F) (W.D.N.Y. May 21, 2013) ("Gordon I"); *Gordon v. Kaleida Health*, 2013 WL 2250579, No. 08-CV-378S(F) (W.D.N.Y. May 21, 2013) ("Gordon II"); *Hinterberger v. Catholic Health System, Inc.*, 2013 WL 2250591, No. 08-CV-380S(F) (W.D.N.Y. May 21, 2013) ("Hinterberger I"); *Hinterberger v. Catholic Health System, Inc.*, 2013 WL 2250603, No. 08-CV-380S(F) (W.D.N.Y. May 21, 2013) ("Hinterberger II").

⁷⁰ *Gordon I*, 2013 WL 2250506; see also *Hinterberger I*, 2013 WL 2250591. Defendants hired the same vendor, approximately one year prior to Plaintiffs' retention of the vendor, to scan boxes of documents and to "objectively code" them. Defendants moved to disqualify Plaintiffs' expert and Plaintiffs' counsel due to this issue. The court declined to disqualify either Plaintiffs' expert or Plaintiffs' counsel, finding, among other reasons, that scanning and objectively coding documents did not constitute "expert" work but was ministerial in nature and that there was no evidence that any confidential information was conveyed to Plaintiffs. *Id.*

⁷¹ *Gordon II*, 2013 WL 2250579 at *1; *Hinterberger II*, 2013 WL 2250603 at *1.

⁷² *Id.*

Due to the dispute over plaintiffs' consulting expert, plaintiffs claimed that defendants refused to meet and confer regarding a predictive coding protocol.⁷³ However, defendants did send a proposed protocol to plaintiffs and indicated that they would also send a list of email custodians.⁷⁴

Plaintiffs "contend[ed] that where a party intends to use predictive coding to assist in the review and production of ESI, it is necessary that the parties negotiate a proposal to guide the use of predictive coding software for the case."⁷⁵ Specifically, "Plaintiffs maintain[ed] Defendants' position excludes Plaintiffs' access to important information regarding Defendant's selection of so-called 'seed set documents' which are used to 'train the computer' in the predictive coding search method."⁷⁶ Defendants claimed "that courts do not order parties in ESI disputes to agree to specific protocols a computer-based review of ESI based on the general rule that ESI production is within the 'sound discretion' of the producing party."⁷⁷ Defendants also stated that the *Da Silva Moore* court "did not direct defendants in that case to provide plaintiffs with the 'seed set documents' defendants intended to use in connection with predictive coding, rather, defendants volunteered to provide such data."⁷⁸ Since defendants indicated their willingness to meet and confer regarding the protocol once the dispute regarding plaintiffs' consulting expert was resolved, and "[b]ased on Defendants' expressed awareness of Defendants' discovery obligations, the court also need not, as Plaintiffs request, remind Defendants of relevant considerations regarding Defendants' use of predictive coding regarding ESI document production obligations."⁷⁹ The court dismissed Plaintiffs' motion without prejudice.⁸⁰

H. *In the Matter of the Search of Information Associated with the Facebook Account Identified by the Username Aaron.Alexis that is Stored at Premises Controlled by Facebook, Inc.* – Judge Facciola

Judge Facciola, one of the forerunners of all judicial e-discovery pioneers, was presented with an opportunity to give his thoughts on TAR in *In the Matter of the Search of Information Associated with the Facebook Account Identified by the Username Aaron.Alexis that is Stored at Premises Controlled by Facebook, Inc.*⁸¹ There, Judge Facciola issued an opinion "explain[ing] the Court's reasons for issuing [a] modified search and seizure warrant," having previously found that the government's request was overbroad.⁸² The court explained its concerns about how information that is not relevant to the criminal case is dealt with, and the modified search warrant required that "[a]ll records and content that the government determines [were] NOT within the scope of the investigation, as described above, must either be returned to Facebook, Inc., or, if copies (physical or electronic), destroyed."⁸³ The court was specifically concerned about information seized that related to third parties.⁸⁴ In discussing the problems with overly broad electronic seizures, the court expounded as follows:

73 *Id.*

74 *Id.*

75 *Id.* At *2.

76 *Id.*

77 *Id.*

78 *Id.*

79 *Id.* at *3 (citations omitted).

80 *Id.*

81 2013 WL 7856600, Case No. 13-MJ-742(JMF) (D.D.C. Nov. 26, 2013).

82 *Facebook*, 2013 WL 7856600 at *1.

83 *Id.* at *3.

84 *Id.* at *4.

Finally, since the 2009 amendment to Rule 41, there has been a sea change in the manner in which computers, which now contain enormous amounts of data, are searched with technology assisted review replacing other forms of searching, including the once thought gold standard of file-by-file and document-by document review. Thus, the premise of the 2009 amendment – that law enforcement had to open every file and folder to search effectively – may simply no longer be true. Indeed, this Court finds it hard to believe that a law enforcement agency of remarkable technical ability such as the FBI is opening every file and folder when it seizes a computer that contains a terabyte of data. The Court cannot imagine that it has the time or personnel to do it, nor see any reason to do it when there are more efficient means to do what its agents have to do. Thus, the boilerplate that has appeared in every search warrant application for as long as law enforcement has been searching computers insisting that the agents must open every file and folder may simply be incorrect and therefore an illegitimate premise for the kind of searching law enforcement will actually do.⁸⁵

The court suggested that in the future, the government might consider using “a special master with authority to hire an independent vendor to use computerized search techniques,” or to have a “search protocol...designed to uncover only the information for which it has probable cause.”⁸⁶ The court concluded that “If the government cannot adopt stricter search parameters in future applications, it may find this Court unwilling to issue any search and seizure warrants for electronic data that ignore the constitutional obligations to avoid ‘general’ electronic warrants that are as offensive to the Fourth Amendment as the searches that led to its enactment.”⁸⁷

I. *Federal Housing Finance Agency v. HSBC North America Holdings, Inc.* – Judge Cote

In *Federal Housing Finance Agency v. HSBC North America Holdings Inc.*,⁸⁸ the court denied a request for reconsideration of a discovery order.⁸⁹ In discussing the request for reconsideration, the court expounded support for predictive coding:

Indeed, at the earliest stages of this discovery process, JPMorgan Chase was permitted, over the objection of FHFA, to produce its documents through the use of predictive coding. The literature that the Court reviewed at that time indicated that predictive coding had a better track record in the production of responsive documents than human review, but that both processes fell well short of identifying for production all of the documents the parties in litigation might wish to see.⁹⁰

⁸⁵ *Id.* at *8.

⁸⁶ *Id.*

⁸⁷ *Id.*

⁸⁸ 2014 WL 584300 2014 WL 584300, Nos. 11 Civ. 6189 (DLC), 11 Civ. 6190 (DLC), 11 Civ. 6193 (DLC), 11 Civ. 6195 (DLC), 11 Civ. 6198 (DLC), 11 Civ. 6200 (DLC), 11 Civ. 6201 (DLC), 11 Civ. 6202 (DLC), 11 Civ. 6203 (DLC), 11 Civ. 6739 (DLC), 11 Civ. 7010 (DLC) (S.D.N.Y. Feb. 14, 2014).

⁸⁹ *Federal Housing Finance Agency*, 2014 WL 584300 at *1.

⁹⁰ *Id.* at *3.

In denying the request to challenge the completeness of document productions, the court stated that “[N]o one could or should expect perfection from the [discovery] process. All that can be legitimately expected is a good faith, diligent commitment to produce all responsive documents uncovered when following the protocols to which the parties have agreed, or which a court has ordered.”⁹¹

VI. EPILOGUE - LESSONS FROM THE JUDICIAL PIONEERS

The opinions of the judicial pioneers discussed above manifest a virtual unanimous consensus of support for technology-assisted review. These judicial trail-blazers embrace technology-assisted review for two main reasons. First, empirical data scientifically establishes that TAR equals or exceeds human manual review in search and production reliability. In addition to Judge Peck, several later judicial pioneers cited to sources that assert that predictive coding is more reliable than key-word searching or manual review.⁹² Second, and perhaps most importantly, TAR reduces the expense of document production, especially in cases involving many gigabytes and/or terabytes of electronically stored information. Other courts have followed Judge Peck’s lead in noting the significant cost savings possible with predictive coding.⁹³ Moreover, it is notable that at least two courageous judicial explorers allowed a party to use predictive coding over the objections of the opposition,⁹⁴ while two other pioneering courts raised the use of predictive coding *sua sponte*.⁹⁵

Da Silva Moore used the phrase “computer-assisted review,” but the majority of the later pioneers used the term “predictive coding.”⁹⁶ As new forms of TAR evolve utilizing more sophisticated artificial intelligence techniques and algorithms, predictive coding may rapidly become as *passé* as keyword searching. If predictive coding is today’s “new black,” it could just as easily become tomorrow’s floppy disk. The important takeaway is not that new technology always fails us in the end, but rather a new generation of technology, that is faster, better and cheaper, will always save us in the end.

Future judicial travelers into the deep space of TAR will need to resolve the very contentious issue of transparency – how much cooperation and sharing of information about a party’s TAR process is enough? Where should a court draw the line between good faith compliance with *The Sedona Conference Cooperation Proclamation*, and violation of the attorney work product doctrine? Specifically, should parties be **required** over its objection to share information about the composition and coding of its seed set, or a subset of the seed set? The *Biomet* court encouraged the recalcitrant party to share non-privileged and relevant documents from the seed set, but found that it “did not have any discretion” to compel such cooperation.⁹⁷ Judge Foschio raised the issue of whether the seed set must be shared in *Gordon II* and *Hinterberger II*, but did not resolve it, ultimately ordering the parties to meet and confer further.⁹⁸ While Judge Peck had before him a producing party

91 *Id.* at *2.

92 *See, e.g., National Day Laborer Organizing Network*, 877 F. Supp. 2d 87; *Chevron*, 2013 WL 1087236; *Federal Housing Finance Agency*, 2014 WL 584300.

93 *Gabriel Technologies Corp.*, 2013 WL 410103; *Chevron*, 2013 WL 1087236.

94 *Global Aerospace Inc.*, 2012 WL 1431215; *Federal Housing Finance Agency*, 2014 WL 584300.

95 *EORHB, Inc.*, 2012 WL 4896670; *Chevron*, 2013 WL 1087236.

96 *Global Aerospace Inc.*, 2012 WL 1431215; *National Day Laborer Organizing Network*, 877 F. Supp. 2d 87; *EORHB, Inc.*, 2012 WL 4896670; *Chevron*, 2013 WL 1087236; *In re Biomet*, 2013 WL 1729682, 2013 WL 6405156; *Gordon II*, 2013 WL 2250579; *Hinterberger II*, 2013 WL 2250603; *Federal Housing Finance Agency*, 2014 WL 584300. *But see Gabriel Technologies Corp.*, 2013 WL 410103, *10 (“computer-assisted, algorithm-driven document review”); *National Day Laborer Organizing Network*, 877 F. Supp. 2d 87, 109 (“‘computer assisted’ or ‘predictive coding’”); *Facebook, Inc.*, 2013 WL 7856600, *8 (“technology assisted review”).

97 *In re Biomet*, 2013 WL 6405156 at *2.

98 *Gordon II*, 2013 WL 2250579 at *3; *Hinterberger II*, 2013 WL 2250603 at *3.

that volunteered transparency, his parting advice to future TAR-nizens is that the more transparency that you are willing to give, the more cooperation from opposing counsel, and the more approval from the court, you will be likely to get. We could aptly name this principle “Peck’s Golden Rule of Transparency.”

As the opinions of the ten judicial pioneers discussed above appear to show, TAR is a true litigation game-changer that is here to stay. Professional responsibility requires understanding and competency in its application on the part of the litigation team – a standard of care best met with, rather than without, competent e-discovery counsel. It remains to be seen whether TAR marches forward with the same force as e-discovery has over the past decade, or whether willful blindness, Luddite-like fear of technology, and/or addiction to unchecked legal bills will create obstacles in the path of TAR’s manifest destiny. We thank all of the judicial pioneers of TAR, past, present and future, for showing us all the way forward.

THE SEDONA CONFERENCE COMMENTARY ON PATENT DAMAGES & REMEDIES*

*A Project of The Sedona Conference Working Group
on Patent Damages & Remedies (WG9)*

<i>Author:</i>	The Sedona Conference	
<i>Editor-in-Chief:</i>	John M. Desmarais	
<i>Executive Editor:</i>	Tamir Packin	
<i>Judicial Review Panel:</i>	Hon. James F. Holderman Hon. Barbara M.G. Lynn	Hon. Susan Illston
<i>Chapter Editors:</i>	Donald R. Banowit Jan M. Conlin Rachel Krevans Edward G. Poplawski	Michael L. Brody Andrea Weiss Jeffries James W. Morando
<i>Contributing Editors:</i>	Gregory S. Arovas Marta Beckwith Peter Gafner Noreen Krall Bruce McFarlane Brian Platt Robert G. Sterne Michael J. Wagner Jennifer Yokoyama	Samuel F. Baxter Ahmed J. Davis R. Eric Hutz John LaBarre Hon. Paul R. Michel (ret.) Christopher J. Renk Barry Ungar Loria Yeadon

Thanks go to all who participated in the dialogue that led to this Commentary.

We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 9. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.*

PREFACE

Welcome to the Public Comment Version of The Sedona Conference Commentary on Patent Damages and Remedies, a project of The Sedona Conference Working Group on Patent Damages and Remedies (WG9). This is one of a series of working group commentaries published by The Sedona Conference, a 501(c)(3) research and educational organization that exists to allow leading jurists, lawyers, experts, academics, and others, at the cutting edge of issues in the areas of antitrust law, complex litigation, and intellectual property rights, to come together – in conferences and mini-think-tanks called Working Groups – and engage in true dialogue, not debate, in an effort to move the law forward in a reasoned and just way.

The mission of WG9, formed in November 2010, is “to create guidelines that will help to clarify and guide the evolution of patent damages and remedies considerations to encourage patent damages and remedies law to remain current with the evolving nature of patents and patent ownership.” The Working Group consists of over sixty active members representing all stakeholders in patent litigation. To develop this Commentary, the Working Group held numerous conference calls over the past several years, and the draft was the focus of dialogue at The Sedona Conference’s 14th Annual Conference on Patent Litigation in Del Mar, CA in October 2013.

The Commentary represents the collective efforts of many individual contributors. On behalf of The Sedona Conference, I thank everyone involved for their time and attention during the drafting and editing process, and in particular: Donald R. Banowitz, Michael L. Brody, Jan M. Conlin, John M. Desmarais, Andrea Weiss Jeffries, Rachel Krevans, James W. Morando, Tamir Packin, and Edward G. Poplawski. The Working Group was also privileged to have the benefit of candid comments by several active district court judges with extensive patent litigation trial experience, including the Honorable James F. Holderman, the Honorable Susan Illston, and the Honorable Barbara M.G. Lynn, who all served as the Judicial Review Panel for this Commentary, as well as the Honorable Cathy Ann Bencivengo, the Honorable James L. Robart, and the Honorable Ronald M. Whyte, who also reviewed and commented on the draft. The statements in this Commentary are solely those of the non-judicial members of the Working Group; they do not represent any judicial endorsement of the recommended practices.

Working Group Series output is first published in draft form and widely distributed for review, critique and comment, including in-depth analysis at Sedona-sponsored conferences. Following this period of peer review, the draft publication is reviewed and revised by the Working Group taking into consideration what is learned during the public comment period. Please send comments to us at info@sedonaconference.org, or fax them to us at 602-258-2499. The Sedona Conference hopes and anticipates that the output of its Working Groups will evolve into authoritative statements of law, both as it is and as it should be.

Craig W. Weinlein
Executive Director
The Sedona Conference
June 2014

FOREWORD

As nine-figure and even ten-figure patent damages jury verdicts become more common, patent damages law has become increasingly important. Even though the forty-year-old *Georgia-Pacific* framework for calculating reasonable royalties remains good law, patent damages law remains one of the most complex, unpredictable, and rapidly evolving areas of the law. Indeed, in many cases the parties' expectations with respect to patent damages often differ by orders of magnitude. This, of course, makes resolving cases short of trial much more difficult. Moreover, even a jury verdict may not add sufficient clarity or certainty to allow the parties to resolve remaining disputes. While a large number of jury verdicts remain undisturbed, many jury verdicts regarding patent damages are being overturned by the Court of Appeals for the Federal Circuit or even by district courts in posttrial rulings.

In this paper, Working Group 9 provides principles and best practices in an effort to add clarity and predictability to the area of patent remedies. Participants and observers of the Working Group included a diverse group of attorneys, including inside counsel for patent holders (including non-practicing entities), inside counsel for practicing entities who often find themselves as defendants in patent litigation, and outside counsel representing both patentees and accused infringers. The Working Group also included expert witnesses who are regularly tasked with writing expert reports assessing patent damages. Members of the federal judiciary also participated as observers to the Working Group.

This paper provides a consensus set of principles and best practices that the Working Group believes will move the law forward in a reasoned and just way. The Working Group began its undertaking by focusing on the statutory mandate that damages should be "adequate to compensate for infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer." With that perspective in mind, the Working Group revisited the *Georgia-Pacific* framework for calculating damages, ultimately recommending a departure from the *Georgia-Pacific* framework of establishing a hypothetical negotiation at the time infringement began in favor of a "retrospective" approach to the hypothetical negotiation. Under the retrospective approach, the hypothetical negotiation takes place at the time of trial and allows for consideration of all relevant facts and circumstances occurring up to the time of trial. The Working Group also provides guidelines and best practices regarding several *Georgia-Pacific* factors, and deals with critical issues including: apportionment; the entire market value rule; whether settlement agreements should be considered in the hypothetical negotiation framework; and the appropriate post-verdict legal and equitable remedies available to patent holders. The Working Group also provides best practices for substantive and procedural damages issues regularly arising before, during, and after trial.

This paper does not attempt to address all the issues that arise in the context of remedies for patent infringement; rather, it puts forth guidelines and best practices that can be applied consistently across cases. With respect to patent damages, the paper focuses on reasonable royalty damages because the Working Group felt that it could make a significant contribution in that area of the law. The paper does not address lost profits damages. The paper also does not address the effects of obligations to license patents on fair, reasonable, and non-discriminatory (FRAND) terms, nor does it deal with enhanced damages, such as those potentially available after a finding of willful patent infringement.

John M. Desmarais
Editor-in-Chief, Chair, Working Group 9 Steering Committee

TABLE OF CONTENTS

I. Background	61
Brief History of Patent Remedies	61
Current State of the Law Regarding the Determination of a Royalty	64
<i>Recent Federal Circuit Cases on Use of the Reasonable Royalty to Calculate Damages</i>	64
<i>Recent District Court Cases on Use of the Reasonable Royalty to Calculate Damages</i>	67
The Practical Realities of the Evolving Royalty Law.....	69
<i>Bifurcation of the Trial</i>	69
<i>Posttrial Rulings on Reasonable Royalty Calculations</i>	71
Posttrial Relief and the Issue of On-Going Royalties	72
<i>Injunctions</i>	72
<i>Alternatives to Injunctions</i>	73
<i>Attorneys' Fees and Fee Shifting</i>	74
II. Principles for the Royalty Paradigm	75
Principle II-1: The reasonable royalty in patent infringement matters should fairly compensate the patent holder for the actual use made by the infringer of the patented invention and should be determined by considering what fully informed and reasonable persons in the position of the patent owner (or owners throughout the period of infringement) and the infringer would agree to at the time of trial as a fair price for the use of the patented invention, from the time of first infringement through the time of trial, taking into account all relevant facts and circumstances occurring before or during that period.	75
Principle II-2: The entire market value of the accused product should only be used as the royalty base for the reasonable royalty determination when the patented aspect(s) of the product is (are) shown to form the basis or substantially all of the basis for consumer demand. The evidence to be considered may include evidence of consumer demand any time prior to trial.	78
Principle II-3: Where the entire market value rule does not apply, it is necessary to apportion the revenue associated with the infringing product between its patented and unpatented features. In so doing, it may be appropriate to consider the smallest saleable unit containing the feature or embodying the patented method for use as the apportioned royalty base. The evidence to be considered in assessing apportionment may relate to any time period prior to trial.	80
Principle II-4: Where the accused product incorporates multiple technologies, once the proper royalty base has been determined, the reasonable royalty rate should reflect the relative contribution of the patented invention as compared to the other technologies incorporated into the royalty base. All technologies incorporated into the royalty base prior to trial should be considered. This approach should help to alleviate the problem often referred to as royalty stacking.	82

Principle II-5: A reasonable royalty must reflect the extent to which, throughout the period of infringement, the patented invention has represented an improvement over available alternatives at the time of infringement, including the prior art. A royalty which over- or under-values the inventive contribution of the patent claim is not reasonable.83

Principle II-6: Three principles apply to the consideration of a non-infringing alternative design, or “design around,” in determining a reasonable royalty, as noted below:84

Principle II-6(a): Evidence of a non-infringing design around that is technically and commercially feasible and available during the damages period is relevant to the reasonable royalty determination. A design around need not actually have been implemented in order to be considered, but must be raised during fact discovery to prevent expense, delay, and prejudice.84

Principle II-6(b): In order to be considered a design around in the first instance, the proposed alternative design must be shown by a preponderance of the evidence not to infringe the asserted claims of the patent(s)-in-suit, and to be an acceptable substitute.84

Principle II-6(c): On a proper showing, the total economic cost of the infringer’s next best available alternative may serve to cap the damages award.84

BEST PRACTICES.....87

Principle II-7: Where the technology claimed in the asserted patent is necessary to practice because (1) it is essential to a *de facto* standard or a standard adopted by a recognized standard setting organization (i.e., standard-essential); (2) a technically feasible non-infringing design around alternative is restricted or prohibited by government regulations or requirements; and/or (3) the technically feasible design around is cost-prohibitive, then the reasonable royalty should exclude any premium the patent may command solely resulting from the adoption of the standard or the governmental/commercial prohibitions on design modification. All standards adopted prior to trial may be considered.87

Principle II-8: The comparison of any proposed comparable license to the hypothetical license should itemize and separately value – to the extent possible – the material ways in which the two differ.89

Principle II-8(a): Evidence of a bare patent license to the patent(s)-in-suit in the same field of use as the accused product/service is generally relevant to the reasonable royalty inquiry and should usually be considered in the determination of the reasonable royalty.91

Principle II-8(b): On a proper showing, evidence of a license that is not a bare patent license to the patent(s)-in-suit in the same field of use, but which does license the patent(s)-in-suit as part of a more comprehensive license, may be relevant to the reasonable royalty inquiry.91

Principle II-8(c): On a proper showing, evidence of a license that does not license the patent(s)-in-suit may be relevant to the reasonable royalty inquiry.92

Principle II-8(d): On a proper showing, license agreements in settlement of litigation that license the patent(s) or technology-in-suit may be relevant to the reasonable royalty determination.92

Principle II-9: Whether a reasonable royalty should be structured as a running royalty or a lump sum should be explicitly considered in the reasonable royalty analysis.	96
III. Pretrial Principles and Best Practices	97
Principle III: If and when the court believes that significant questions may exist as to the admissibility of certain damages theories or determinations, then in the appropriate case, the court should consider conducting a hearing after the parties exchange damages contentions to determine: (1) if the parties' damages theories are legally cognizable; (2) if the damages evidence is reliable, relevant and/or admissible; and (3) other disputes relating to damages.	97
BEST PRACTICES.....	98
IV. Trial Principles and Best Practices	104
Principle IV-1: Courts can assist in streamlining the presentation of damages evidence at trial to ensure that: (1) damages theories are tied to the specific facts of the case, and that damages experts use reliable methodologies; (2) the entire market value rule is applied only when appropriate; and (3) the comparability of license agreements is rigorously addressed.....	104
Principle IV-2: A significant amount of trial time should be dedicated to the damages portion of a patent case.	106
BEST PRACTICES.....	106
Principle IV-3: Bifurcation of a patent damages trial from a patent liability trial may be appropriate.	106
Principle IV-3(a): In cases involving a single defendant and a single patent, bifurcation of damages may not be appropriate given the relative lack of complexity in the case, potential overlap in proof on various liability and damages issues, and the risk of prejudice to the patentee if infringement continues unabated throughout the time that it takes to try both phases of a case.....	106
Principle IV-3(b): In cases involving multiple defendants, multiple patents and multiple accused products, or those involving particularly complex damages theories, or those in which the courts order timed trials, bifurcation of damages may be appropriate to avoid juror confusion and unnecessary expense.	106
Principle IV-3(c): If a court decides to completely bifurcate liability discovery and trial from damages discovery and trial, it should consider also allowing time for an appeal to the Federal Circuit between trials.....	106
BEST PRACTICES.....	109
Principle IV-4: In a typical case, a willfulness allegation should not itself dictate a bifurcation of damages from liability. To the extent possible, where a case is bifurcated, and willfulness is tried after liability is determined, it is preferable to have a staged trial with the same jury rather than different juries.	109
Principle IV-5: Jury instructions that are tailored to the case will be more suitable than model jury instructions.....	110
Principle IV-6: Jury verdict forms that are tailored to the case will be more suitable than general verdict forms. Thus, in most cases, the verdict form should ask the jury to determine an amount of damages adequate to compensate for the infringement, on a per patent/per claim basis. Also, special verdict forms may be preferable in cases involving ongoing damages.	111

V. Posttrial Principles and Best Practices112

 A. Injunctions112

 Principle V-1: A patent is a property right and the patentee usually is irreparably harmed if the right to exclude is not enforced.112

 BEST PRACTICES.....116

 B. Alternatives to Injunctions117

 Principle V-2: If an injunction is not available, then ongoing royalties may be available.117

 Principle V-3: An ongoing royalty should fairly compensate the patent holder for the ongoing use made by the infringer of the patented invention and should be determined by considering what fully informed and reasonable persons in the position of the patent owner (or owners throughout the period of infringement) and the infringer would agree to at the time of trial as a fair price for the license, from the time of trial through the expiration of the patent, taking into account all relevant facts and circumstances occurring before or during that period.119

 BEST PRACTICES.....119

 C. Attorneys’ Fees and Fee Shifting120

 Principle V-4: Pursuant to 35 U.S.C. § 285, where a party to a patent lawsuit improperly initiates or maintains one or more claims or defenses, an award of attorneys’ fees is presumptively appropriate. Attorneys’ fees may be assessed against the party and/or its counsel as circumstances warrant.120

 Principle V-5: A claim or defense is improper at whatever point in time it becomes the case that:121

 1. The patentee and/or its counsel had actual knowledge that (or were willfully blind as to the fact that) the asserted claim is either not valid, not infringed, or not enforceable;.....121

 2. The accused infringer and/or its counsel had actual knowledge that (or were willfully blind as to the fact that), contrary to an asserted defense, the asserted claim is valid, infringed, or enforceable;.....121

 3. In the case of a party to a patent lawsuit, a reasonable person in the position of the patentee and advised by competent counsel would understand that the pursuit of a claim or defense is without merit; or121

 4. In the case of the party’s counsel, competent counsel in the position of the party’s counsel would understand that the pursuit of a claim or defense is without merit.121

 Principle V-6: Indicators that a reasonable person in the position of a party and/or the party’s counsel would know that the initiation or maintenance of a claim or defense is improper include, but are not limited to, the following:122

 1. The claim or defense rests on a construction of a claim limitation that (a) was explicitly disclaimed during prosecution or in the specification, or (b) is objectively inconsistent with the plain meaning of the limitation, and the plain meaning of the limitation is not disclaimed elsewhere in the intrinsic record;122

2. The party or original patentee (where the original patentee is not a party) or its counsel previously had made a statement about the patent to a court, the Patent and Trademark Office, or another administrative body that cannot reasonably be reconciled with the initiation or the maintenance of a claim or defense;122

3. There is evidence (a) which establishes as a matter of law that a claim or defense is objectively baseless, and (b) which, after the initiation of a lawsuit, is actually called to the party’s attention through discovery, or, prior to the initiation of a lawsuit, was obtained, or through the exercise of reasonable diligence could have been obtained, from the public record or from witnesses under the control of the party; or122

4. There is a reasonable basis to believe that a case was brought for the purpose of obtaining a settlement of a meritless claim for materially less than the likely cost of litigation.....123

VI. Conclusion123

I. BACKGROUND

Brief History of Patent Remedies

The Constitution grants Congress the power “to promote the Progress of Science and the Useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.” Congress has exercised that power by granting courts the authority to award compensatory damages for infringement and injunctions prohibiting future infringement.

Patent infringement damages have their current statutory basis in Section 284 of the Patent Act, which states that, upon a finding of infringement, “the court shall award the claimant damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer”¹ In practice, damages awards today typically are based on two general forms of damages: lost profits of the patentee; or a reasonable royalty based on either an established royalty or the framework of a hypothetical negotiation.

Early Supreme Court case law divided along two lines: the first interpreted the then existing patent damages statutes² to require either actual harm to the patentee or an established royalty for the patented technology in order to award damages. For example, the Court in one early case held the patentee was entitled to only nominal damages upon a finding of infringement, because the patentee did not prove any established royalty and “[t]here was no question . . . of damages arising from lost sales, or injurious competition, for no machines had been manufactured and put on the market by the patentee.”³ The Court further offered a definition of what constitutes an established reasonable royalty, noting that “it must be paid or secured *before* the infringement complained of; it must be paid by such a number of persons as to indicate a general acquiescence in its reasonableness by those who have occasion to use the invention; and it must be uniform at the places where the licenses are issued.”⁴

The next year, the Court followed this same approach in *Coupe v. Royer*, pointing to the lower court’s error in permitting a jury to award infringement damages without evidence of a true licensing fee or impairment of the patentee’s market in any way. As such, since there were “no damages of any kind . . . the lower court should have instructed the jury . . . to find nominal damages only.”⁵

The second line of case law focused instead on the actual value of the invention to determine an appropriate measure of damages, without regard to any established royalty rate or actual harm to the patentee. For example, the Court affirmed a damages award following a jury instruction permitting examination of “the utility and advantage[s] of the

1 35 U.S.C. § 284.

2 Section 14 of the Patent Act of 1836 provided, in part: “[I]t shall be in the power of the court to render judgment for any sum above the amount found by such verdict as the actual damages sustained by the plaintiff, not exceeding three times the amount thereof, according to the circumstances of the case, with costs”

In 1870, Congress amended the statute to provide that for suits in equity, “the claimant [complainant] shall be entitled to recover, in addition to the profits to be accounted for by the defendant, the damages the complainant has sustained thereby, and the court shall assess the same or cause the same to be assessed under its direction, and the court shall have the same powers to increase the same in its discretion that are given by this act to increase the damages found by verdicts in actions upon the case”

3 *Rude v. Westcott*, 130 U.S. 152, 167 (1889).

4 *Id.* at 165 (emphasis added).

5 *Coupe v. Royer*, 155 U.S. 565, 583 (1895).

invention over the old modes or devices that had been used for working out similar results,” despite a lack of evidence of either harm to the patentee or an established royalty rate.⁶ The Court affirmed this view in 1871, characterizing the damages question as requiring a determination of what advantages the defendant derived from using the patented invention over simply using other processes that were legally available for public use, which would have allowed him to obtain “an equally beneficial result.”⁷

Similarly, an 1853 case noted that an inventor of an improvement to a mill should not be permitted to claim damages arising from lost profits on the entire mill; rather, damages should be measured based solely on the use of the inventor’s improvement to the mill.⁸ The decision in *Westinghouse Elec. & Mfg. Co. v. Wagner Elec. & Mfg. Co.* elaborated on this principle, stating that where a patentee’s invention “only created a part of the profits, he is only entitled to recover that part of the net gains.”⁹ Further, the patentee in such a case must produce evidence of apportionment of the profits between the patented features and the remaining features to distinguish between the patentee’s damages and the defendant’s rightful profits.¹⁰ However, the *Westinghouse* Court also stated that “when it is impossible to make a mathematical or approximate apportionment[,] . . . [o]n established principles of equity, and on the plainest principles of justice, the guilty trustee cannot take advantage of his own wrong[.]” and the patentee is entitled to all of the infringer’s profits.¹¹

Many viewed *Westinghouse* as enabling patentees to recover excessive damages for infringement in too many cases. Just three years later, in *Dowagiac Mfg. v. Minn. Moline Plow*, the Court retreated from the expansive *Westinghouse* decision, holding that “[i]n the absence of [an established] royalty, and in the absence of proof of lost sales or injury by competition,” the patentee bore the burden of proving a reasonable royalty.¹² The Court then remanded the case, giving the patentee the opportunity to show the invention’s actual value by “proving what would have been a reasonable royalty, considering the nature of the invention, its utility and advantages, and the extent of the use involved.”¹³

As noted above, the current approach to damages derives from Section 284 of the Patent Act, and most typically requires a determination of the patentee’s lost profits or what constitutes a reasonable royalty. While an established royalty may form the basis for a reasonable royalty under the current law, it is rare for such an established royalty to exist because the particular invention or technology at issue may not have been licensed out to other entities in the same factual context. As such, courts have come to use the hypothetical negotiation framework and rely on the numerous factors enumerated in *Georgia-Pacific Corp. v. U.S. Plywood Corp.* to determine what constitutes a reasonable royalty in any given case. The factors include:

6 *Suffolk Co. v. Hayden*, 70 U.S. 315, 320 (1865).

7 *Mowrey v. Whitney*, 81 U.S. 620, 651 (1871).

8 *See Seymour v. McCormick*, 57 U.S. 480, 489 (1853). This view is a precursor to today’s entire market value rule, which applies to both lost profits damages and reasonable royalty damages. This rule applies when the invention is one element of a product sold, and states that a patentee is only entitled to damages based on the invention itself (not the entire product), unless the patented element is the basis – or a substantial basis – for demand of the entire product. *See infra* Chapter II, Principle II-2.

9 *Westinghouse Elec. & Mfg. Co. v. Wagner Elec. & Mfg. Co.*, 225 U.S. 604, 615 (1912).

10 *See id.* Again, this is consistent with the entire market value rule and current principles of apportionment in cases awarding damages based on lost profits. *See infra* Chapter II, Principle II-2.

11 *Id.* at 620.

12 *Dowagiac Mfg. Co. v. Minn. Moline Plow Co.*, 235 U.S. 641, 649 (1915). However, the Court did not expressly overrule the statements made in *Westinghouse*, which remained good law until the Patent Act amendments in 1946, which eliminated the patentee’s right to disgorgement of all of the defendant’s profits under *Westinghouse*. *See* Erick S. Lee, *Historical Perspectives on Reasonable Royalty Patent Damages and Current Congressional Efforts for Reform*, 13 UCLA J.L. & TECH. 1, 8 (2009) (noting the 1946 amendments’ abrogation of the use of an infringer’s profits as a basis for measuring damages).

13 *Dowagiac*, 235 U.S. at 648.

1. royalties received by the patentee for licensing the patent-in-suit, proving or tending to prove an established royalty;
2. rates paid by the licensee for the use of comparable patents;
3. nature and scope of the license, as exclusive or non-exclusive, or as restricted or non-restricted in terms of territory or with respect to whom the manufactured product may be sold;
4. licensor's established policy and marketing program to maintain his patent monopoly by not licensing or by granting licenses under special conditions designed to preserve that monopoly;
5. commercial relationship between the licensor and licensee, such as whether they are competitors or whether they are inventor and promoter;
6. effect of selling the patented invention in promoting sales of other products of the licensee, and existing value of the invention to the licensor as a generator of sales of his non-patented items;
7. duration of the patent and the term of the license;
8. established profitability of the product made under the patent, its commercial success, and its current popularity;
9. utility and advantages of the patent property over the old modes or devices, if any, that had been used for working out similar results;
10. nature of the patented invention, character of the commercial embodiment of it as owned and produced by the licensor, and benefits to users of the invention;
11. extent to which the infringer has made use of the invention, and any evidence probative of the value of that use;
12. portion of the profit or selling price that may be customary in the industry to allow for the use of the invention or similar inventions;
13. portion of the realizable profit that should be credited to the invention as distinguished from non-patented elements, the manufacturing process, business risks, or significant features or improvements added by the infringer;
14. opinion testimony of qualified experts;
15. amount that a prudent licensor and a prudent licensee would have agreed upon at the time the infringement began, if both had been reasonably and voluntarily trying to reach an agreement.¹⁴

While the hypothetical negotiation and the *Georgia-Pacific* factors remain a well-accepted framework for calculating reasonable royalty damages today, the *Georgia-Pacific* factors leave significant room for interpretation. Because of that, in recent years the Federal Circuit has issued many decisions evaluating appropriate methods and considerations when calculating infringement damages based on a reasonable royalty. Cases addressing these issues have become common in recent years and courts are finding a need to ensure that such awards are based on accepted methodologies and sufficient evidence.

¹⁴ *Georgia-Pacific Corp. v. U.S. Plywood Corp.*, 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970).

Current State of the Law Regarding the Determination of a Royalty¹⁵

Recent Federal Circuit Cases on Use of the Reasonable Royalty to Calculate Damages

The Federal Circuit stated in *Lucent*, in 2009, that several approaches may be used to calculate a reasonable royalty, including: (1) the analytical method, which calculates damages based on the infringer's anticipated profit from sales of the infringing product; and (2) the "more common" hypothetical negotiation approach contemplated in *Georgia-Pacific*.¹⁶

In *Lucent*, lump sum damages of roughly \$358 million were awarded to Lucent for Microsoft's indirect infringement.¹⁷ The Federal Circuit vacated the damages award on the ground that it was not supported by substantial evidence.

At trial, both parties advocated for the hypothetical negotiation approach. Accordingly, the Federal Circuit reviewed the damages award by applying the *Georgia-Pacific* framework. At trial, Lucent asked for damages based on a running royalty, while Microsoft argued that any damages were represented by a lump sum royalty payment of \$6.5 million. Because the jury verdict awarded a lump sum, paid-in-full royalty of about \$358 million, on appeal the Federal Circuit evaluated whether substantial evidence supported the jury's implicit finding that at the time of the hypothetical negotiation, Microsoft would have agreed to a lump sum, paid-in-full royalty of about \$358 million.¹⁸

In addressing *Georgia-Pacific* factor 2 – "[t]he rates paid by the licensee for the use of other patents comparable to the patent in suit" – the Federal Circuit held that the licenses presented by Lucent at trial were for other groups of patents, and were created from contexts far different from a license negotiation tailored to the patent-in-suit.¹⁹ The Federal Circuit, applying *Georgia-Pacific* factors 10 and 13, found that the infringing feature contained in Microsoft Outlook is but a tiny feature of one part of a much larger software program, and that it was inconceivable to conclude, based on the record below, that the use of that small feature constituted a substantial portion of the value of Microsoft Outlook. Accordingly, the Federal Circuit held that *Georgia-Pacific* factors 10 and 13 provided little support for the jury's lump sum damages award.²⁰

In analyzing *Georgia-Pacific* factor 11 – "[t]he extent to which the infringer has made use of the invention; and any evidence probative of the value of that use" – the Federal Circuit relied on the "book of wisdom"²¹ to reject Microsoft's argument that information about consumers' use of the infringing feature was irrelevant because it postdated the time of the hypothetical negotiation. In allowing such post-hypothetical negotiation evidence, the Federal Circuit explained that "neither precedent nor economic logic requires us to ignore information about how often a patented invention has been used by infringers. Nor could they since frequency of expected use and predicted value are

15 This paper does not address the impact of Fair, Reasonable, and Non-Discriminatory (FRAND) obligations on the royalty analysis or recent adaptations of the *Georgia-Pacific* factors to account for those obligations.

16 *Lucent Techs., Inc. v. Gateway, Inc.*, 580 F.3d 1301, 1324 (Fed. Cir. 2009).

17 *See id.* at 1325.

18 *Id.*

19 *Id.* (citing *Georgia*, 318 F. Supp. at 1120).

20 *See Lucent*, 580 F.3d at 1332–33.

21 While the hypothetical negotiation constructively takes place at or before the time the infringement began, the methodology is flexible and it may accommodate a court's inquiry into "events and facts that occurred [after the hypothetical negotiation] that could not have been known to or predicted by the hypothesized negotiators." *Fromson v. Western Litho-Plate & Supply Co.*, 853 F.2d 1568, 1575 (Fed. Cir. 1988). This is known as "the book of wisdom," or evidence arising after the infringement began. *See Sinclair Refining Co. v. Jenkins Petroleum Process Co.*, 289 U.S. 689, 698 (1933) (Cardozo, J.) ("[I]f years have gone by before the evidence is offered, [e]xperience is then available to correct uncertain prophecy. Here is a book of wisdom that courts may not neglect.").

related.”²² The Federal Circuit held that *Georgia-Pacific* factor 11 did not support the jury verdict because “the evidence of record is conspicuously devoid of any data about how often consumers use the patented date-picker invention.”²³

The *Lucent* court noted that while the determination of the reasonable royalty “must relate to the time infringement occurred, and not be an after-the-fact assessment,”²⁴ evidence of subsequent events “can, under appropriate circumstances, be helpful to the jury and the court in assessing whether a royalty is reasonable.”²⁵ In rejecting Microsoft’s argument that Lucent should not be permitted to rely on evidence concerning consumer use of the patented feature due to its generation post-negotiation, the Federal Circuit stated that such information may aid the hypothetical negotiation calculation, since it provides information that parties would have had to estimate if done at the time of negotiation.²⁶

In evaluating the decision below, the Federal Circuit also held that, to the extent the jury relied on an entire market value rule calculation to arrive at the lump sum damages amount, that award was not supported by substantial evidence and was against the clear weight of the evidence, for two reasons:

First, Lucent had failed to show that the patented invention provided “the basis – or even a substantial basis – of the consumer demand for Outlook,” a necessary condition for application of the entire market value rule. Rather, the evidence demonstrated that the infringing date-picker tool was “but a very small component of a much larger software program” and that the vast majority of Outlook’s features did not infringe. “Indeed Lucent’s damages expert conceded that there was no ‘evidence that anybody anywhere at any time ever bought Outlook . . . because it had a date picker.’”²⁷

Second, Lucent’s damages expert used the wrong approach in explaining how the entire market value rule should be applied. Initially, the expert sought to apply a royalty rate of 1% to a royalty base consisting of the price of the entire computer loaded with the infringing software. After the district court excluded this testimony, the expert changed the royalty base to the price of the software alone but increased the royalty rate to 8% in order to obtain the same damages number. As “there was no evidence that Microsoft had ever agreed to pay an 8% royalty on an analogous patent,” the Federal Circuit held that the expert’s approach “d[id] not comport with the purpose of damages law or the entire market value rule.”²⁸

In 2010, the Federal Circuit took a similar view in *ResQNet.com v. Lansa, Inc.*, vacating a damages award because it “relied on speculative and unreliable evidence divorced from proof of economic harm linked to the claimed invention.”²⁹ The patentee’s damages expert based his royalty opinion on the *Georgia-Pacific* framework, assessing a “starting point” for the hypothetical negotiation based on the first *Georgia-Pacific* factor – royalties received by the patentee from existing licenses. But the first *Georgia-Pacific* factor focuses on “licensing of the patents-in-suit, proving or tending to prove an established royalty.”³⁰ In his

22 *Lucent*, 580 F.3d at 1333.

23 *Id.* at 1334.

24 *Unisplay, SA v. Am. Elec. Sign Co.*, 69 F.3d 512, 518 (Fed. Cir. 1995).

25 *Lucent*, 580 F.3d at 1333–34.

26 *See id.* at 1334. However, the Federal Circuit ultimately determined that Lucent could not rely on the book of wisdom approach, as the “record [was] conspicuously devoid of any data about how often consumers use[d] the patented . . . invention.” *Id.*

27 *Id.* at 1337–38.

28 *Id.* at 1338.

29 *ResQNet.com, Inc. v. Lansa, Inc.*, 594 F.3d 860, 868 (Fed. Cir. 2010).

30 *Georgia-Pacific*, 318 F. Supp. at 1120.

explanation to the jury, the patentee's damages expert referred to license agreements bearing no relation to the invention at issue; these licenses furnished source code and services, and had no "discernible link to the claimed technology"; yet the expert relied solely on these agreements to find support for an "unjustified" royalty rate in the double-digits.³¹ As such, the district court erred in adopting the expert's proposed royalty rate and failing to make an effort to link the licenses to the patented technology.³² In *dicta*, the Federal Circuit "observe[d] as well that the most reliable license in this record arose out of litigation."³³ On remand, the Federal Circuit further instructed the district court that it "may also consider the panoply 'of events and facts that occurred thereafter and could not have been known to or predicted by the hypothesized negotiators.'"³⁴

In *Wordtech Sys., Inc. v. Integrated Networks Solutions, Inc.*, the Federal Circuit found error in the jury's reliance on non-comparable license agreements in awarding a lump sum reasonable royalty to the patentee.³⁵ The Federal Circuit held that the patentee's evidence of thirteen prior licenses was insufficient to support the jury's award, as only 2 of those 13 were lump sum agreements. The award was an approximate average of those two lump sum licenses, but the Federal Circuit held that even those licenses were insufficient because "they provide[d] no basis for comparison with [the] infringing sales."³⁶ Specifically, "[n]either license describe[d] how the parties calculated each lump sum, the licensee's intended products, or how many products each licensee expected to produce."³⁷

Three years ago, the Federal Circuit disavowed the use of the "25 percent rule of thumb" as a "fundamentally flawed tool for determining a baseline royalty rate in a hypothetical negotiation . . . because it fails to tie a reasonable royalty base to the facts of the case at issue."³⁸ There, the patentee's damages expert presented no evidence that a 25%/75% split was standard practice in beginning Uniloc's license negotiations, nor did Uniloc attempt to show that the patented invention's contribution to the accused products justified a 25% royalty. Further, the Federal Circuit rejected the expert's use of Microsoft's total revenue as a "check" on the reasonableness of the proposed royalty rate, noting that precedent does not support use of the entire market value rule in the case of minor patent improvements even if the asserted royalty rate is low enough.³⁹

In *LaserDynamics, Inc. v. Quanta Computer, Inc.*, the Federal Circuit revisited the entire market value rule and the admissibility of settlement agreements.⁴⁰ The Federal Circuit made clear that to satisfy the entire market value rule, the patented feature must be the motivating factor for the purchase of the product, not merely a "required" or "important" feature. The Federal Circuit also discussed the use of settlement agreements as evidence to establish the amount of a reasonable royalty, stating that the "propriety of using prior settlement agreements to prove the amount of a reasonable royalty is questionable."⁴¹ The Federal Circuit reasoned that settlement agreements are of questionable propriety due to the difference between the circumstances of litigation as compared to the legal fiction of a hypothetical negotiation resulting in an agreement between willing licensees and licensors. In *LaserDynamics*, a particular settlement agreement LaserDynamics sought to

31 *ResQNet.com*, 594 F.3d at 870.

32 *See id.* at 873.

33 *Id.* at 872.

34 *Id.* (quoting *Fromson*, 853 F.2d at 1575).

35 *See Wordtech Sys., Inc. v. Integrated Networks Solutions, Inc.*, 609 F.3d 1308 (Fed. Cir. 2010).

36 *Id.* at 1320.

37 *Id.*

38 *Uniloc USA, Inc. v. Microsoft Corp.*, 632 F.3d 1292, 1315 (Fed. Cir. 2011).

39 *See id.* at 1320.

40 *See LaserDynamics, Inc. v. Quanta Computer, Inc.*, 694 F.3d 51 (Fed. Cir. 2012).

41 *Id.* at 77.

introduce was excluded. The license amount in the settlement agreement was many times more than the amount other licensees negotiated outside of litigation, likely because the defendant was facing trial and sanctions. The licenses granted outside of litigation were “far more reliable indicators of what willing parties would agree to in a hypothetical negotiation.”⁴² It was therefore improper for LaserDynamics’ expert to selectively rely on the license amount from the settlement, while ignoring the licenses voluntarily negotiated outside of litigation; the royalty rate arrived at by LaserDynamics’ expert was “untethered from the patented technology at issue and the many licenses thereto, and, as such, was arbitrary and speculative.”⁴³

As indicated above, case law has interpreted Section 284 to permit damages based on either a reasonable royalty or lost profits. The Federal Circuit’s explicit recognition in *Lucent* that the statute permits multiple and varying approaches for the determination of a reasonable royalty, and that the approach of *Georgia-Pacific* is only one permissible approach,⁴⁴ paves the way for the consideration of new alternative approaches, which is one of the primary focuses of this Commentary.

Recent District Court Cases on Use of the Reasonable Royalty to Calculate Damages

In light of recent Federal Circuit case law, including those cases discussed above, district courts are taking a variety of approaches with respect to: (1) the entire market value rule; (2) alternatives to the hypothetical negotiation between patentee and defendant; (3) the issue of whether licenses are sufficiently “comparable”; and (4) the admissibility of settlement agreements.

With regard to the entire market value rule, district courts have taken different approaches in their application of the rule. For example, the Southern District of New York excluded an expert’s testimony on the entire market value rule because, in its view, the expert had applied the wrong standard.⁴⁵ Rather than opine that the patented feature was “the” basis for customer demand, he had opined that it was “a substantial basis for demand.”⁴⁶ In contrast, the Eastern District of Texas allowed testimony on the entire market value, despite the fact that it was undisputed that the patented feature did not provide the basis for customer demand, because between 13 and 16 comparable licenses also provided for a royalty based on the entire value of the licensed products.⁴⁷

There is uncertainty and variation in application of the “hypothetical negotiation” as the paradigm for determining the appropriate reasonable royalty. The Southern District of California denied a defendant’s motion to exclude a patentee’s damages expert from testifying about “real-world” negotiations.⁴⁸ Under the expert’s theory, the parties would enter the negotiation with their “respective walk away approaches” and ultimately “meet in the middle.”⁴⁹ The defendant argued that the expert’s approach was not appropriately grounded in the facts of the case, disputing the patentee’s expert use of several methodologies in support of the expert’s conclusions.⁵⁰ The court, however, rejected the

42 *Id.* at 78.

43 *Id.* at 81.

44 See also *WhitServe, LLC v. Computer Packages, Inc.*, 694 F.3d 10 (Fed. Cir. 2012) (“We do not require that witnesses use any or all of the *Georgia-Pacific* factors when testifying about damages in patent cases.”).

45 See *Inventio AG v. Otis Elevator Co.*, No. 06 Civ. 5377, 2011 WL 3359705, at *6 (S.D.N.Y. June 23, 2011).

46 *Id.*

47 See *Mondis Tech. v. LG Electronics, Inc.*, 2:07-CV-565, 2011 WL 2417367, at *2–6 (E.D. Tex. June 14, 2011).

48 *Lucent Techs., Inc. v. Microsoft Corp.*, No. 07-cv-2000, 2011 WL 7664416, at *16 (S.D. Cal. June 16, 2011).

49 *Id.*

50 See *id.*

defendant's argument, noting that "[I]tigators routinely adopt several approaches for calculating a reasonable royalty" and concluding the business realities negotiation theory employed by the patentee's expert was based on reliable principles and methods.⁵¹

A number of district courts have also been receptive to the idea of using evidence generated subsequent to the timing of the hypothetical negotiation. For example, the court in *LecTec Corp. v. Chattem, Inc.*, declined to exclude the testimony of a damages expert regarding events subsequent to the date of the hypothetical negotiation, holding that criticism of the book of wisdom approach is "better directed to weight rather than admissibility."⁵² Similarly, the District of Delaware has rejected a defendant's argument that the patentee's expert gave "subsequent events too much weight in his royalty calculation" in denying the defendant's motion to exclude expert testimony.⁵³

District courts have also taken divergent approaches when deciding whether licenses are sufficiently "comparable." Some courts have looked with disfavor on the use of "industry" licenses where the licenses encompass far more than the technology at issue in the case.⁵⁴ Courts have also taken expert witnesses to task where the experts failed to take into account existing licenses to the claimed technology that would have been "appropriate as touchstones for determining the appropriate royalty rate."⁵⁵ Courts seem to be increasingly scrutinizing expert testimony about why the licensed technology is comparable to the technology claimed in the patent-in-suit, and whether the circumstances of prior licenses are comparable to the circumstances between the parties.⁵⁶

Other courts have taken a more relaxed approach to "comparability." For example, in a case involving a patented stent, the Eastern District of Texas decided that licenses were sufficiently comparable when they related to a drug delivered by the stent, or to a method of drug delivery that was similar to the method employed by the patented product.⁵⁷ Similarly, the District of Delaware upheld a jury determination that a patentee receive over \$9 million for the infringement of three of its patents even though Microsoft had paid just \$8 million to license the patentee's entire portfolio.⁵⁸ The court reasoned that the discrepancy was, in part, the result of "the substantial intangible benefits that stem from being endorsed by Microsoft."⁵⁹ The "true value" of the Microsoft license could therefore reasonably have exceeded \$8 million.⁶⁰

The Federal Circuit has recognized the difficulty of identifying "comparable" licenses, realizing that upon close inspection, few, if any, "real world licenses introduced at trial [arise] from circumstances *identical* to those presumed to prevail in the hypothetical royalty negotiation."⁶¹ Laboring, perhaps, under *Uniloc's* emphatic reiteration of *Lucent's* exacting standard, the Eastern District of Virginia recently lamented that "[a]ll five licenses [it considered in a case] contained various restrictive limitations as well as the rights to use

51 *Id.*

52 *LecTec Corp. v. Chattem, Inc.*, No. 5:08-CV-130, slip op. at 4 (E.D. Tex. Feb. 1, 2011).

53 *St. Clair Intellectual Prop. Consultants, Inc. v. Canon, Inc.*, No. 03-241, 2004 U.S. Dist. LEXIS 19475, at *9 (D. Del. Sept. 28, 2004).

54 *IP Innovation LLC v. Red Hat, Inc.*, 705 F. Supp. 2d 687, 691 (E.D. Tex. 2010).

55 *See, e.g., id.*

56 *See, e.g., Wordtech Sys.*, 609 F.3d at 1320; *Fenner Investments, Ltd. v. Hewlett-Packard Co.*, No. 6:08-cv-273, 2010 U.S. Dist. LEXIS 101384, at *5 (E.D. Tex. Apr. 16, 2010); *Ricoh Co., Ltd., v. Quanta Computer, Inc.*, 06-cv-462-bbc, 2010 U.S. Dist. LEXIS 27301 at *26–27 (W.D. Wis. Mar. 23, 2010).

57 *See Saffran v. Johnson & Johnson*, No. 2:07-CV-451 (TJW), 2011 U.S. Dist. LEXIS 34858, at *31–32 (E.D. Tex. Mar. 31, 2011).

58 *See Finjan Software, Ltd. v. Secure Computing Corp.*, C.A. No. 06-369 (GMS), 2009 U.S. Dist. LEXIS 72825, at *41 (D. Del. Aug. 18, 2009), *rev'd on other grounds*, 626 F.3d 1197 (Fed. Cir. 2010).

59 *Id.* at *39.

60 *Id.*

61 *Lucent*, 580 F.3d at 1329 (emphasis added).

the patents. Each license also reflected the result of different perceived litigation strength[s] and weaknesses based on litigation developments,” including verdicts, a hung jury and the early-state settlements.⁶² The court concluded:

[Because] each of the five licenses reflects unique considerations which defy quantification . . . [we] cannot envision a reasonable, reliable way to use those five licenses to arrive at an ongoing royalty

[Furthermore,] the [c]ourt harbors serious doubt as to whether it has any authority to incorporate the various intangible provisions included in the prior licenses.⁶³

Just as district courts have taken different approaches to determine whether licenses are comparable, or whether the entire market value is appropriately in play, they have also taken divergent approaches regarding the admissibility of litigation-induced settlement licenses. Some courts have denied the admission of settlement agreements because of their view that the potential for prejudice and jury confusion substantially outweighs the licenses’ probative value.⁶⁴ There also is the risk that permitting use of litigation settlement agreements will result in a more complicated trial in which a lot of time and energy will be devoted to evidence relating to the circumstances that caused the litigation settlement agreement. Other courts, however, have admitted settlement agreements on a case-by-case basis when they: (1) are the only sufficiently comparable license(s) to the patent-in-suit; or (2) closely resemble comparable, non-litigation-induced licenses.⁶⁵

The Practical Realities of the Evolving Royalty Law

Bifurcation of the Trial

Because of the increase in the complexity of damages theories, the need for flexibility given the varied district court approaches to damages issues, and the length of time required to try the damages portion of a patent case, a significant consideration going forward is whether damages should be tried together with liability issues, or bifurcated and tried at a later date.

Under the Federal Rules, bifurcation is proper for “convenience, to avoid prejudice, or to expedite and economize.”⁶⁶ District courts have broad discretion in determining whether to bifurcate.⁶⁷ The party seeking bifurcation bears the burden of demonstrating that it is proper given the facts of the case.⁶⁸

62 *ePlus, Inc. v. Lawson Software, Inc.*, No. 3:09-cv-620, 2011 U.S. Dist. LEXIS 54957, at *46 (E.D. Va. May 23, 2011).

63 *Id.* at *46, 47 n.9.

64 *See, e.g., Fenner Inv. Ltd. v. Hewlett Packard Co.*, 2010 U.S. Dist. LEXIS 41514, at *10 (E.D. Tex. Apr. 28, 2010); *Cornell Univ. v. Hewlett-Packard Co.*, 2008 U.S. Dist. LEXIS 39343, at *13–14 (N.D.N.Y. May 8, 2008).

65 *See ReedHycalog UK, Ltd. v. Diamond Innovations, Inc.*, 727 F. Supp. 2d 543, 546 (E.D. Tex. 2010); *see also ResQNet.com*, 594 F.3d at 872.

66 FED. R. CIV. P. 42(b).

67 *See, e.g., Robert Bosch, LLC v. Pylon Mfg. Corp.*, 719 F.3d 1305, 1319–20 (Fed. Cir. 2013) (making clear that “district courts, in their discretion, may bifurcate willfulness and damages issues from liability issues in any given case” and that “[d]istrict court judges, of course, are best positioned to make [the determination to bifurcate issues] on a case-by-case basis”); *Shum v. Intel Corp.*, 499 F.3d 1272, 1276 (Fed. Cir. 2007); *York v. AT&T*, 95 F.3d 948, 957 (10th Cir. 1996); *Idzjotic v. Pa. Railroad Co.*, 456 F.2d 1228, 1230 (3d Cir. 1972).

68 *See, e.g., Brown v. Toscano*, 630 F. Supp. 2d 1342, 1345 (S.D. Fla. 2008); *Windsor Indus., Inc. v. Pro-Team, Inc.*, 87 F. Supp. 2d 1129, 1131 (D. Colo. 2000).

Advantages to bifurcation – in particular in multi-defendant or multi-patent/multi-accused product cases – include potential cost savings and efficiencies. In certain cases, it makes little sense to incur the costs associated with fact and expert damages discovery, which can be quite substantial, unless and until a determination is made on the extent to which any defendant is liable.

However, bifurcation may result in duplicative efforts where evidence on liability issues overlaps with the proof required to support damages theories. Further, should damages be determined by a different jury, bifurcation may put either, or both, patentee and defendant at a strategic disadvantage. A patentee loses the benefit of the jury having full knowledge of all of the proof of the defendant's wrongdoing when it is determining the royalty to be awarded.⁶⁹ Similarly, an accused infringer may be at a disadvantage where the damages jury has no knowledge of its non-infringement and/or invalidity arguments.

Before filing suit, a plaintiff-patentee should be cognizant of any local rules or practices regarding bifurcation. Certain districts have local rules regarding bifurcation and individual judges may have “a preference [on bifurcation] based on past experience” from which “they rarely deviate.”⁷⁰ For example, for several years one judge in the District of Delaware adhered to a standard patent scheduling order under which damages and willfulness were bifurcated from liability “unless good cause is shown.”⁷¹ The judge's rationale was that “discovery disputes related to document production on damages and the *Daubert* motion practice related to damages experts are a drain on scarce judicial resources.”⁷² In each instance in which no liability is found, the time spent mediating discovery disputes or making damages *Daubert* determinations is utterly wasted. This judge also believed that parties are likely to settle after liability has been found to avoid an unpredictable damages award.⁷³ Settlement discussions after a liability determination are believed to “give the parties – those with the most expertise in the market – the first opportunity to translate the [court's] final legal decision on liability into practical commercial consequences.”⁷⁴

By contrast, in the Northern District of Georgia, bifurcation is unlikely. The district's local rules state that “[t]here shall be a rebuttable presumption against the bifurcation of damages from liability issues in patent cases for purposes of either discovery or trial.”⁷⁵ Similarly, individual judges from the District of Utah, Northern District of Texas, Southern District of Florida, Northern District of Illinois, and the Northern District of Indiana have also stated a presumption against bifurcation.⁷⁶ In the view of these judges, bifurcation results in duplicative discovery, witnesses and evidence, and simply delays final

69 35 U.S.C. § 284.

70 *Bifurcation Ruling Highlights Divergent Approaches to Patent Case Management*, THE DOCKET REPORT (Aug. 31, 2009, 10:41 AM), <http://docketreport.blogspot.com/2009/08/bifurcation-ruling-highlights-divergent.html>.

71 See Hon. Sue Robinson, *Standard Patent Scheduling Order*, para. 2(a) (revised 12/4/09), available at <https://web.archive.org/web/20100527104709/http://www.ded.uscourts.gov/SLR/Forms/SchedOrder-Patent.pdf> (“The issues of willfulness and damages shall be bifurcated for purposes of discovery and trial, unless good cause is shown otherwise.”). Judge Robinson's recently revised standard patent scheduling order no longer bifurcates willfulness and damages issues from liability issues. See Hon. Sue Robinson, *Standard Patent Scheduling Order* (revised 3/24/14), available at <http://www.ded.uscourts.gov/sites/default/files/Chambers/SLR/Forms/Sched-Order-Patent-03-24-14.pdf>.

72 *Robert Bosch, LLC v. Pylon Mfg. Corp.*, Civ. No. 08-542-SLR, 2009 WL 2742750, at *1 (D. Del. Aug. 26, 2009); *Dutch Branch of Streamserve Dev., AB v. Extracore Software LLC*, Civ. No. 08-343-SLR, 2009 U.S. Dist. LEXIS 76006, at *2 (D. Del. Aug. 26, 2009).

73 See *Robert Bosch*, 2009 WL 2742750, at *1.

74 *Id.*; *Dutch Branch*, 2009 U.S. Dist. LEXIS 76006, at *3.

75 Patent L.R. 5 (N.D. Ga.), available at <http://www.gand.uscourts.gov/pdf/NDGARulesPatent.pdf>.

76 See *Lutron Elecs. Co., Inc. v. Creston Elecs., Inc.*, No. 2:09-cv-707 DB, 2010 U.S. Dist. LEXIS 49623 (D. Utah May 19, 2010); *Nielsen v. Alcon, Inc.*, Civ. A. No. 3:08-CV-02239-B, 2010 U.S. Dist. LEXIS 26804 (N.D. Tex. Mar. 22, 2010); *Baratta v. Homeland Housewares, LLC*, No. 05-cv-60187 (S.D. Fla. Oct. 27, 2008); *DSM Desotech Inc. v. 3D Sys. Corp.*, 2008 U.S. Dist. LEXIS 87473 (N.D. Ill. Oct. 28, 2008); *BASF Catalysts LLC v. Aristo, Inc.*, No. 2:07-cv-222, 2009 U.S. Dist. LEXIS 16263 (N.D. Ind. Mar. 2, 2009).

resolution.⁷⁷ They believe that the requisite level of complexity that warrants bifurcation simply does not exist when there is only one patent-in-suit, where the technology is straightforward and easy to understand, or where the court will only have to grapple with issues common to many (or all) patent cases, including claim construction, an assessment of the prior art, or the resolution of inventorship disputes.⁷⁸

Litigants should be aware that bifurcation may have a significant effect on the admissibility of evidence. For example, the District of Maryland has bifurcated a trial where a defendant wanted to raise an “advice of counsel” defense to avoid a willfulness finding and treble damages, but did not want to waive attorney-client privilege with respect to liability issues.⁷⁹ Similarly, the District of Delaware has acknowledged that evidence of a previous verdict in favor of the plaintiff-patentee – and evidence of prior licensing agreements – was relevant to patentee’s damages theories⁸⁰ but nevertheless excluded the evidence. Given that liability and damages were being tried together, the court concluded that the evidence posed a substantial danger of unfair prejudice to defendant.⁸¹

Bifurcation may also allow for the admission of damages evidence that is better tailored to the extent to which a defendant is liable. If the extent of liability has not yet been determined, parties may present damages evidence that far exceeds – or grossly underestimates – the true scope of the injury to the patentee. At least one district court has suggested that bifurcation would allow the parties to present the jury with more accurate damage estimates, noting “several instances in which damages evidence will be admissible only if certain factual predicates are established.”⁸² By establishing facts related to liability and the scope of the injury through a bifurcated trial, damages experts would be limited in their estimates and would present a more accurate picture to the jury.

Should a litigant desire to bifurcate damages from liability, in a district where bifurcation is not the norm, a motion to bifurcate should be brought early in the case. If parties have already completed extensive discovery related to damages, any benefits of reaching the liability issues faster will already be limited. Similarly, if parties have constructed their litigation strategy around the assumption of a single trial, it is not productive to require them to redevelop their plans after the preliminary stages of litigation. Thus, bifurcation may be less appealing to the court once discovery is underway.

Posttrial Rulings on Reasonable Royalty Calculations

Posttrial challenges to reasonable royalty calculations can be difficult. For example, the Western District of Wisconsin has upheld a patentee’s expert’s argument that worldwide royalty rates should be adjusted upwards for application in the United States because patent enforcement is much more common in this country.⁸³ The trial judge did not question that analysis because the jury awarded less than the full measure of damages the expert recommended.⁸⁴ Because the jury adopted a lower figure, the judge determined that even if the patentee had failed to support its view, he would not say that there was “no rational connection between the award and the evidence.”⁸⁵

77 See *Baratta*, 05-cv-60187, slip op. at 9; *Nielsen*, 2010 U.S. Dist. LEXIS 26804, at *5.

78 See *id.*

79 *Mike’s Train House, Inc. v. Broadway Ltd. Imports, LLC*, Civ. No. JKB-09-2657, 2011 U.S. Dist. LEXIS 27550 (D. Md. Mar. 17, 2011).

80 See *St. Clair Intellectual Prop. Consultants, Inc. v. Fuji Photo Film Co., Ltd.*, 674 F. Supp. 2d 555, 559 (D. Del. 2009).

81 See *id.*

82 *United States Gypsum Co. v. Lafarge N. Am. Inc.*, 670 F. Supp. 2d 737, 747 (N.D. Ill. 2009).

83 See *Ricoh*, 2010 U.S. Dist. LEXIS 27301, at *28–29.

84 See *id.*

85 *Id.*

Similarly, the District of Minnesota, despite being “initially troubled” by a jury’s damages verdict it declared “certainly generous,” has upheld a damages determination because it had “sufficient basis in the evidence at trial” and did not “reflect a miscarriage of justice.”⁸⁶ Faced with evidence that the damages may exceed the cost of a non-infringing alternative, the court reasoned that “a reasonable jury . . . could have disregarded this proposed non-infringing alternative.”⁸⁷

And, the Northern District of Ohio has upheld a jury’s damages award that was outside the range established by the parties’ experts.⁸⁸ In that case, both parties’ experts agreed that 4% was a reasonable royalty for a hypothetical licensing agreement between two willing parties.⁸⁹ The plaintiff’s expert, however, emphasized that the plaintiff licensor was “not anxious to grant a license,” and the jury decided on damages exceeding a 4% reasonable royalty.⁹⁰ The court upheld the award, concluding that “when supported by the evidence, a jury may rightfully award damages . . . in excess of any amount advocated by either party.”⁹¹

Posttrial Relief and the Issue of On-Going Royalties

Injunctions

As noted above, the Constitution grants Congress the power “to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”⁹²

Congress has exercised that power, and the Patent Act expressly provides for the granting of a permanent injunction to a successful patentee: “The several courts having jurisdiction of cases under this title may grant injunctions in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable.”⁹³

In *eBay Inc. v. MercExchange, LLC*, the United States Supreme Court took up the issue of the standard that should govern when injunctions are issued in patent cases.⁹⁴ The Court rejected the Federal Circuit’s “general rule that courts will issue permanent injunctions against patent infringement absent exceptional circumstances,” and held that the patentee must satisfy the same four-factor test applied in other injunction contexts by showing: (1) irreparable injury; (2) that the remedies available at law are inadequate to compensate for that injury; (3) that the balance of hardships between the plaintiff and defendant favors an injunction; and (4) the public interest would not be disserved by issuance of an injunction.⁹⁵

86 *Spectrallytics, Inc. v. Cordis Corp.*, 650 F. Supp. 2d 900, 907–08 (D. Minn. 2009), *aff’d in part, remanded on other grounds*, 2011 U.S. App. LEXIS 11981 (Fed. Cir. June 13, 2011).

87 *Id.*

88 *See Bendix Commercial Vehicle, Sys., LLC v. Haldex Brake Prod. Corp.*, No. 1:09 CV 176, 2011 U.S. Dist. LEXIS 138, at *6–7 (N.D. Ohio Jan. 3, 2011).

89 *See id.* at *4.

90 *Id.* at *6–7.

91 *Id.* at *5; *Cf., e.g., Lucent*, 580 F.3d at 1332 (refusing to uphold jury’s damages in part because the jury did not choose “a damages award somewhere between maximum and minimum lump-sum amounts advocated by the opposing parties.”); *Fuji Photo Film Co. v. Jazz Photo Corp.*, 394 F.3d 1368, 1378 (Fed. Cir. 2005) (“[T]he jury is not bound to accept a rate proffered by one party’s expert but rather may choose an intermediate royalty rate.”).

92 U.S. CONST. art. 1, § 8.

93 35 U.S.C. § 283.

94 *See eBay Inc. v. MercExchange, LLC*, 547 U.S. 388 (2006).

95 *Id.* at 391, 393–94.

The majority opinion, however, did not provide any guidance with respect to the weight, if any, that should be given to the previously accepted concept that patents are a property right, which generally should be protected by the right to exclude. This led to two concurring opinions, addressed in Chapter V, which, directly or indirectly, address that question.^{96 97}

Of course, in some instances injunctions remain appropriate. For example, in *Robert Bosch, LLC v. Pylon Manufacturing Corp.*,⁹⁸ the Federal Circuit held that it was an abuse of discretion for the court to decline to award injunctive relief where: (1) the parties were direct competitors; (2) there was a loss of market share and potential customers; and (3) due to financial problems, the infringer might not be able to satisfy a monetary judgment.⁹⁹ The International Trade Commission also continues to grant injunctions, as it does not have the power to award damages and is not bound by the *eBay* factors.

Alternatives to Injunctions

Prior to the Supreme Court's ruling in *eBay*, courts routinely granted injunctions to successful patentees; therefore, there rarely existed a need to determine what remedy was appropriate for post-judgment infringement. Post-*eBay*, determinations regarding ongoing infringement absent an injunction have become important.

Where an injunction is not granted, courts can simply do nothing and await any future suit for further infringement. This approach, however, undoubtedly presents efficiency concerns for the parties and the courts. As such, the issue has arisen whether courts can determine forward damages for ongoing infringement in the same suit. The Federal Circuit has held that “[u]nder some circumstances, awarding an ongoing royalty for patent infringement in lieu of an injunction may be appropriate.”¹⁰⁰ For example, in *ActiveVideo Networks, Inc. v. Verizon Communications, Inc.*, the Federal Circuit concluded that the district court erred in finding money damages inadequate to compensate for the infringement, as the patent holder had engaged in extensive licensing and licensing efforts, had solicited the defendant for a license over a long period of time preceding and during litigation, and there was no direct competition between plaintiff and defendant.¹⁰¹ The Federal Circuit concluded that plaintiff was entitled to an ongoing royalty: “ActiveVideo’s loss of revenue due to Verizon’s infringement can be adequately remedied by an ongoing royalty from Verizon for each of its subscribers. This is what ActiveVideo has sought from Verizon since 2004, and based on the infringement determinations ActiveVideo is certainly entitled to it.”¹⁰²

The Federal Circuit has also held that “[t]here is a fundamental difference, however, between a reasonable royalty for pre-verdict infringement and damages for post-verdict infringement.”¹⁰³ As a result of the Federal Circuit’s limited guidance, various mechanisms for dealing with ongoing royalties in lieu of a permanent injunction have been utilized by district courts, with no common approach having yet been adopted.¹⁰⁴ For

96 See *infra*, Chapter V.A.

97 *Id.* at 394–95 (Roberts, J., concurring); *id.* at 395–97 (Kennedy, J., concurring); see also *Apple, Inc. v. Motorola, Inc.*, 869 F. Supp. 2d 901 (N.D. Ill. June 22, 2012), *rev’d*, 2014 WL 1646435 (Fed. Cir. April 25, 2014).

98 *Robert Bosch, LLC v. Pylon Mfg. Corp.*, 659 F.3d 1142 (Fed. Cir. 2011).

99 See *id.* at 1152–55.

100 *Paice LLC v. Toyota Motor Corp.*, 504 F.3d 1293, 1314 (Fed. Cir. 2007).

101 See *ActiveVideo Networks, Inc. v. Verizon Commc’ns, Inc.*, 694 F.3d 1312, 1339–40 (Fed. Cir. 2012).

102 *Id.* at 1340.

103 *Amado v. Microsoft Corp.*, 517 F.3d 1353, 1361 (Fed. Cir. 2008).

104 See, e.g., *Sovereign Software LLC v. J.C. Penney Corp., Inc.*, 899 F. Supp. 2d 574, 589–90 (E.D. Tex. 2012) (applying a 2.5x enhancement to the jury’s implied royalty rate for ongoing royalties).

instance, may parties ask the jury to determine a fully paid up lump sum to account for future infringement, or must the issue of ongoing infringement absent an injunction be dealt with via an ongoing running royalty? If the jury is permitted to award a lump sum, how should such a lump sum properly be determined?

*Attorneys' Fees and Fee Shifting*¹⁰⁵

The guidelines provided in Chapter V are intended to give greater clarity as to which litigation practices warrant shifting fees and costs pursuant to 35 U.S.C. § 285.¹⁰⁶ When there is clarity around the practices that are unacceptable, these practices appear less attractive to litigants, and consequently, the victims of such practices will be more readily made whole when claims and defenses are nonetheless improperly pursued.

The perceived need for enhanced clarity in this domain arises from the economics of patent litigation. The inherent complexity of a patent case necessarily makes its prosecution or defense a costly undertaking. As a result, there is an ever-present opportunity for both plaintiffs and defendants to arbitrage the cost of litigation into a settlement that is inconsistent with the merits of the claim or defense. Thus, for example, the AIPLA's biennial survey of litigation costs reports that the 2011 median cost through appeal of a patent lawsuit involving \$1 – \$25 million was \$2.5 million. For a patent lawsuit involving more than \$25 million, the 2011 median cost through appeal was \$5.5 million.¹⁰⁷

Absent a meaningful fee shifting remedy, a litigant faced with an adversary's meritless claims or defenses has three options, all of them bad:

1. capitulate, and pay a negotiated ransom to avoid the cost of abusive litigation;
2. fight through an adversary's abusive conduct, and pay the price of abusive litigation directly in unnecessary attorneys' fees; or
3. retaliate in kind, thereby escalating the level of pointless costs by imposing them in both directions.

The only way to shortcut this spiral of futility is to remove the financial incentives that make abusive pursuit of meritless litigation a rational option. That necessarily means shifting the cost of abusively imposed litigation expense to the instigator.

The mechanism for fee shifting is explicit in the statute's recognition that "[t]he court in exceptional cases may award attorney fees to the prevailing party."¹⁰⁸ Such fee shifting, however, is rarely employed. In April 2014, the U.S. Supreme Court lowered the "exceptional" case standard for prevailing parties to collect attorneys' fees from the Federal Circuit's former "objectively and subjectively baseless" standard to one that covers litigation practices that "stand[] out from others with respect to the substantive strength of a party's litigating position (considering both the governing law and facts of the case) or the unreasonable manner in which the case was litigated."¹⁰⁹

¹⁰⁵ This paper does not address willful patent infringement or the potential enhanced damages and attorneys' fees recoverable after such a finding.

¹⁰⁶ See *infra*, Chapter V.C.

¹⁰⁷ See AM. INTELLECTUAL PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY 2011 (2011).

¹⁰⁸ 35 U.S.C. § 285.

¹⁰⁹ See *Octane Fitness, LLC v. Icon Health & Fitness, Inc.*, 134 S.Ct. 1749, 1756 (2014).

Proposed legislation has been recently introduced to address these issues, such as the House bill entitled “Saving High-Tech Innovators from Egregious Legal Disputes Act,” or “SHIELD,” which proposed a fee-shifting provision that would apply to any party “alleging the infringement of [a] patent” in the software or computer hardware field.¹¹⁰ Most recently, for example, the House Judiciary Committee approved H.R. 3309, the “Innovation Act” introduced by Chairman Bob Goodlatte (R-Va.), that includes an attorney fee-shifting provision for patent cases, which would award reasonable fees and expenses to a prevailing party “unless the court finds that the position and conduct of the nonprevailing party or parties were reasonably justified in law and fact or that special circumstances (such as severe economic hardship to a named inventor) make an award unjust.”¹¹¹

The Working Group believes that providing more concrete guidance as to what behavior is not reasonably justified – thereby attaching greater potential monetary risk to its pursuit – will make such behavior less tactically beneficial and, as a result, less common. At a minimum, it is hoped that such guidance will give a meaningful remedy to the victims of such conduct.

II. PRINCIPLES FOR THE ROYALTY PARADIGM

Principle II-1: The reasonable royalty in patent infringement matters should fairly compensate the patent holder for the actual use made by the infringer of the patented invention and should be determined by considering what fully informed and reasonable persons in the position of the patent owner (or owners throughout the period of infringement) and the infringer would agree to at the time of trial as a fair price for the use of the patented invention, from the time of first infringement through the time of trial, taking into account all relevant facts and circumstances occurring before or during that period.

Comment

The Working Group discussed three approaches to the overall reasonable royalty paradigm:

1. The status quo – a hypothetical negotiation at the time of first infringement using only facts available at that time, except for certain future facts that may be taken into account under the “book of wisdom” principle;¹¹²
2. A strict “Prospective Only” model – a hypothetical negotiation at the time of first infringement, using only facts known at that time, and eliminating altogether the book of wisdom exception; and
3. A new “Retrospective” Model – taking into account not only facts available at the time of first infringement but all facts available through the time of trial, eliminating any need for the book of wisdom exception.

110 Saving High-Tech Innovators from Egregious Legal Disputes Act of 2012, H.R. 6245, 112th Cong. (2012) (“Notwithstanding section 285, in an action disputing the validity or alleging the infringement of a computer hardware or software patent, upon making a determination that the party alleging the infringement of the patent did not have a reasonable likelihood of succeeding, the court may award the recovery of full costs to the prevailing party, including reasonable attorney’s fees, other than the United States.”).

111 Innovation Act, H.R. 3309, 113th Cong. (2013); *see also* Patent Abuse Reduction Act of 2013, S. 1013, 113th Cong. (2013).

112 For a description of the “book of wisdom” principle, *see supra* note 21.

The Working Group has determined that the third approach is the most consistent with both the statutory damages provision and sound economic principles. Key reasons for this decision are as follows:

- Due to the amorphous nature of the book of wisdom principle, the status quo approach is seriously flawed. As practitioners appreciate, there are few references to the book of wisdom in the case law, and of those, none provide clear guidance as to the nature and extent to which future facts and circumstances may be taken into account when assessing the appropriate reasonable royalty. The parties and their experts, therefore, cannot reasonably predict what facts they will be permitted to rely upon at trial, and the district court has little help when making its decision. As a result, parties on both sides attempt to “cherry-pick” from the future facts favorable to their case, and omit mention of any future developments that are unfavorable. There is no legal or economic principle that justifies this result. Moreover, the (largely) prospective nature of the status quo approach tends to be applied exclusively to the assessment of the appropriate royalty rate, and not at all to the assessment of the appropriate base to which the royalty rate is to be applied. That is, once the royalty rate of the hypothetical negotiation is determined, it is applied to the actual numbers of units sold from the time of first infringement through trial. These actual units clearly encompass facts not known to the parties at the time of the negotiation, and in many cases, facts not even predictable. Thus, there is a mix-and-match aspect to the status quo approach that is arguably inconsistent.
- The strict prospective model was analyzed extensively. The advantage of this model is that it provides predictability and is consistent with the status quo principle of determining the reasonable royalty at the time of first infringement. That is, if the hypothetical negotiation is placed at the time of first infringement, facts not then available to the parties should not be considered. Allowing consideration of such facts to enter into the negotiation (via the book of wisdom or otherwise) is a clear departure from, and perhaps an erosion of, the principle. While appealing, the Working Group recognized that a rigid prospective model may be at odds with the overriding mandate of the damages statute: to adequately compensate the patent owner for the use made of the invention by the infringer. For example, after the date of first infringement, a key fact may change so as to cause the infringement to be much more detrimental to the patent owner or so as to cause the patent rights to be far more valuable than appeared to be the case earlier. In such instances, the prospective only model would not yield damages adequate to compensate the patent owner for the infringement. Conversely, if after the date of first infringement, a key fact changed such that the patent rights were less valuable to the owner and the infringer, ignoring that fact results in overcompensation to the patent owner and an unfair penalty to the infringer. It was this potential for unfair results that was the catalyst for the creation of the book of wisdom exception to the prospective model in the first instance.
- The Working Group believes that the Retrospective Model is the most economically sound approach that both accomplishes the goals of the patent damages statute and also is consistent with the economic principles governing patent valuation. Taking all facts known through the time of trial into account eliminates the potential for unfairness in the prospective model without

introducing the cherry-picking and uncertainty that the book of wisdom imported into that model. Moving the hypothetical negotiation later in time, however, to a time at or near the time of trial has potential infirmities as well. Specifically, it could lead to a higher (and potentially unfair) royalty due to what are commonly known as “lock in” effects.¹¹³ As discussed below, the Working Group endorsement of the Retrospective Model incorporates a methodology to avoid this potential problem. With that methodology in place, the Retrospective Model can achieve the full purposes of the statute – adequate compensation to the patent owner – without unfairness to either party.

The Working Group’s Retrospective Model states:

The royalty shall be determined by considering what (a) fully informed and (b) reasonable persons (c) in the position of the patent owner or owners throughout the period of infringement and the infringer would agree to at the time of trial as a fair price for the license, taking into account (d) all relevant facts and circumstances occurring before or during that period.

Detailed application of the hypothetical negotiation is discussed in succeeding sections of this report, but the general principles are set out here:

- a. “Fully informed” means a licensor and licensee who both know all relevant facts available to them, not just the facts actually known to the individuals at the plaintiff and defendant entities at the relevant times.
- b. “Reasonable persons” means hypothetical negotiators in the place of the actual plaintiff and defendant, applying an objective, not a subjective, determination. The hypothetical negotiators must be assumed to behave like reasonable business people would under the circumstances.
- c. “[What] the patent owner or owners throughout the period of infringement and the infringer would agree to” means that the hypothetical reasonable person will bargain in light of the actual circumstances of the patent owner(s) and the defendant(s), and the actual context of the market for the invention and products or processes using the invention at the relevant times.
- d. “All relevant facts and circumstances” include the factors listed in *Georgia-Pacific* (to the extent relevant) and any other factors relevant to the particular case being litigated, for example:
 - (i) the relative bargaining power of the patent owner and the accused infringer throughout the period of infringement, including their positions in the market for the accused processes and/or products;
 - (ii) the importance of the accused product or process to the business of the defendant(s);

113 The “lock in” effects are drawing significant attention in the context of fair, reasonable, and non-discriminatory (FRAND) licensing because “one of the primary purposes of the [F]RAND commitment is to avoid patent hold-up, which occurs when the holder of a standard-essential patent demands excess royalties after standard implementers are already locked into using the standard.” *In re Innovatio IP Ventures, LLC Patent Litig.*, No. 1:11-CV-9308, 2013 WL 5593609, at *8-*9 (N.D. Ill. Oct. 3, 2013); see *Microsoft Corp. v. Motorola, Inc.*, No. C10-1823JLR, 2013 WL 2111217, at *10 (W.D. Wash., Apr. 25, 2013); see also, *infra*, Chapter II, Principle II-7 and accompanying commentary.

- (iii) the actual financial position of the patent owner(s) and the accused infringer(s) throughout the period of infringement;
- (iv) the terms of comparable licenses, if any;
- (v) industry practices in license structures and amounts, for example, whether it is industry practice to pay lump sum royalties, or to pay a set amount per unit rather than a rate-times-a-base price;
- (vi) if the appropriate license structure is a rate-times-a-base, what is the appropriate base? That is, how much of the value of the accused product or process is attributable to the claimed invention?;
- (vii) past practice of patent owner(s) and defendant(s) in license structure and amounts;
- (viii) what licenses has the patent owner already given on this technology and will this hypothetical license impair any of them? Are there most favored nations clauses in existing licenses that must be respected? Has the patent owner already given an exclusive license in the field in which the defendant is practicing the invention?;
- (ix) royalty stacking, if any;
- (x) non-infringing alternative design, or “design around,” possibilities, including their cost, technical and commercial feasibility, and time to develop;
- (xi) how the royalty under the hypothetical license would fit into an appropriate cost and profit structure for the defendant(s).

If any of the relevant factors changed materially during the period from the time of first infringement to the time of trial, two different royalty rates might apply – one to the period before the material change, and one to the period after – rather than one rate for the entire period.

Principle II-2: The entire market value of the accused product should only be used as the royalty base for the reasonable royalty determination when the patented aspect(s) of the product is (are) shown to form the basis or substantially all of the basis for consumer demand. The evidence to be considered may include evidence of consumer demand any time prior to trial.

Comment

The entire market value rule (EMVR) allows patent infringement damages to be based on the full market value of the product or process sold in certain circumstances. It originated as a doctrine applicable only to a lost profits analysis but in recent years also has been applied to the reasonable royalty analysis.

The principle articulated above allows for application of the EMVR in the reasonable royalty context, just as it applies in the lost profits context,¹¹⁴ but at the same time avoids the concerns that have been lodged against using the EMVR in the reasonable royalty context, such as those expressed by the FTC, which has commented:

Courts should eliminate the entire market value rule and the question of whether the patented feature was the “basis for customer demand” from the determination of the appropriate base in a reasonable royalty damages calculation. It is irrelevant and it risks injecting significant confusion that threatens to produce inaccurate awards.¹¹⁵

Because many infringement cases involve accused products or processes having features unrelated to the patent-in-suit or which are successful in part for reasons unrelated to the patented feature or method, there is consensus among the Working Group that in the reasonable royalty context, a royalty may be applied to the entire market value of an accused product only in circumstances where the patented feature or method is the basis or substantially all of the basis for demand of the product.¹¹⁶ In cases involving accused products with many (in some cases, as many as hundreds or thousands) of components and inventions, it is unlikely that any one patented feature or method will provide the basis or substantially all of the basis for the demand for the product.

In determining whether a royalty may be applied to the entire market value of an accused product, courts, experts, and parties must exercise diligence to ensure that the appropriate question is asked: Is the patented feature the sole basis or substantially all of the basis for customer demand? For example, it is not sufficient merely to establish, whether through expert or documentary evidence, that “but for” that feature or method a consumer would not purchase the product. In most circumstances involving complex products, there are several “but for” features and/or methods without which there would be no consumer demand. If there is more than one “but for” feature or method, then a royalty should not be applied to the entire market value of an accused product.

In making this determination, courts, experts, and parties also must guard against the tendency to assume away the basic functionality of an accused product or process. Consider the following example: the asserted patent claims a new method of security for routers. The accused infringer’s internal documentation states that consumers would not buy the accused routers if they were not secure. The patent holder conducts a survey asking whether security is important in the decision to purchase a router, and whether security was the basis for the purchase. The respondents universally responded that security is their top priority in deciding which router to purchase, and that they purchased the accused routers because they were secure. For EMVR purposes, the internal documentation and the survey are focused incorrectly – they remove from consideration whether the device in question has all of the various features that provide for basic router functionality in the first place. The proper threshold EMVR question is whether security is the feature – among all

114 See, e.g., *Rite-Hite Corp. v. Kelley Co.*, 56 F.3d 1538, 1549 (Fed. Cir. 1995) (*en banc*) (“When a patentee seeks damages on unpatented components sold with a patented apparatus, courts have applied a formulation known as the ‘entire market value rule’ to determine whether such components should be included in the damage computation, whether for reasonable royalty purposes, . . . or for lost profits purposes . . .”).

115 FED. TRADE COMM’N, THE EVOLVING IP MARKETPLACE: ALIGNING PATENT NOTICE AND REMEDIES WITH COMPETITION, at 25, 211 (2011), available at <http://www.ftc.gov/os/2011/03/110307patentreport.pdf> [hereinafter *FTC Report*].

116 Although not discussed in the text, the entire market value rule also allows for damages based on the entire market value of an accused product where: (i) the infringing and non-infringing components are sold together so that they constitute a functional unit or are parts of a complete machine or single assembly of parts; or (ii) the infringing and non-infringing components are analogous to a single functioning unit. See *Rite-Hite*, 56 F.3d at 1549–50 (Fed. Cir. 1995) (*en banc*); *Cornell Univ. v. Hewlett-Packard Co.*, 609 F. Supp. 2d 279, 286–87 (N.D.N.Y. 2009) (Rader, J., sitting by designation).

product features including the routing features – that drives consumer demand. Security may be very important, but presumably, a consumer would not buy a secure router that did not properly forward data from one network to another.

Moreover, it is important to distinguish the claimed invention from a general characterization of the feature to which the invention pertains. Taking the above example, it is important to determine whether the particular security method claimed in the asserted patent was the basis for consumer demand, as opposed to other non-infringing security methods. That is, did consumers purchase the accused router because of the particular security method employed to secure the router's data transmissions? This might be the case, or consumers may have purchased the accused router simply because it was secure, irrespective of the particular security method. Furthermore, the asserted security method may be a subset of numerous other technologies encompassed in accused router's security technology, which may suggest that no one patent is responsible for the demand for the security feature.

Notwithstanding the foregoing, it may be appropriate under certain circumstances to use the entire market value of the accused product even in the absence of satisfying the EMVR test articulated as Principle II-2. For example, there may be an established practice whereby the entire market value of a product is used as the royalty base in bona fide license agreements licensing the patented technology for use in products analogous to the accused products. In such a case, it may be appropriate to use the entire market value of the accused products as the royalty base and the concomitant royalty rate provided for in the bona fide comparable licenses, rather than to apportion the entire market value of the accused products and adjust the industry license rate(s) or risk rendering the bona fide licenses non-comparable.¹¹⁷ Such an approach would be supported by *Georgia-Pacific* factor 1 (the royalties received by the patent owner for the licensing of the patent-in-suit, proving or tending to prove an established royalty).

Ultimately, it is the responsibility of the court to address whether the EMVR is satisfied prior to trial. If the patent holder cannot establish that the patented invention is the sole basis or substantially all of the basis for consumer demand (or whose only evidence of this is of the type of improper evidence discussed above), then the royalty base must be properly apportioned, and the patentee should not be permitted to rely upon the EMVR or present a royalty based upon the entire market value of the accused products or services.

Principle II-3: Where the entire market value rule does not apply, it is necessary to apportion the revenue associated with the infringing product between its patented and unpatented features. In so doing, it may be appropriate to consider the smallest saleable unit containing the feature or embodying the patented method for use as the apportioned royalty base. The evidence to be considered in assessing apportionment may relate to any time period prior to trial.

Comment

This principle recognizes that where the EMVR does not apply due to a finding that the patented feature does not form the requisite basis for customer demand, it is

¹¹⁷ See, e.g., *Mondis*, 2011 WL 2417367 at *2-3 (denying motion to strike use of EMVR notwithstanding failure to meet the EMVR test because comparable licenses used the entire market value as the royalty base).

important, when determining the base, to allocate the portion of the value of the product that is the result of the patented feature. A reasonable royalty should be based on the incremental value the patented invention adds to the overall product or service; i.e., “[t]he portion of the realizable profit that should be credited to the invention as distinguished from non-patented elements, the manufacturing process, business risks, or significant features or improvements added by the infringer.”¹¹⁸ Other product features, as well as other factors (e.g., product marketing, the goodwill associated with the entity making or selling the product or service, the sales team, the sales model, and the availability and type of additional products, services or support available from the entity) may contribute to the revenue of a particular product or service.

Ideally, apportionment isolates the patented feature from other features and factors that contribute to the patented product or service’s value. There are a number of potential starting points for determining the appropriate royalty base. In some circumstances, the best starting point will be the smallest saleable unit containing or utilizing the patented invention.¹¹⁹ In other circumstances, the most appropriate starting point will be the incremental difference in value between a product containing or utilizing the patented invention and a similar product that does not include or utilize the patented invention. In yet other circumstances, the right starting point will be the amount paid for the component that includes or utilizes the patented invention.

Irrespective of which starting point is used, the objective of the court, experts, and parties involved is to apportion the royalty base as closely as possible to reflect the incremental value attributable to the patented invention. Stated differently, the starting point may not be the end point of the analysis. Taking the smallest saleable unit as an example, it may be necessary to further apportion below the smallest saleable unit starting point.¹²⁰

Additionally, the fact that patent claims may be expressed in different formats, and of different scope, does not allow the patent owner to avoid apportionment.¹²¹ A patent, for example, may be claimed as a combination of a known apparatus or method in combination with additional elements or steps that constitute the improvement, or it may be claimed in a format where the improvement is explicitly recited (known as a “Jepson” claim). As the FTC put it:

Another artificial construct for identifying the base that courts should reject is always to equate it with the device recited in the infringed claim. In many cases, there will be an easy correspondence between the inventive feature, the device recited in the infringed claim, and the appropriate base. In other cases, the correspondence will not be so clear. For example, a software invention for rendering video images can be recited in a claim covering video software, or in a claim covering a standard personal computer running the video software . . . “[T]he

118 *Garretson v. Clark*, 111 U.S. 120, 121 (1884) (noting that the patentee “must in every case give evidence tending to separate or apportion the defendant’s profits and the patentee’s damages between the patented feature and the unpatented features”); *Georgia-Pacific*, 318 F. Supp. at 1120.

119 *See, e.g., Cornell Univ.*, 609 F. Supp. 2d at 288.

120 *See, e.g., Rembrandt Social Media, LP v. Facebook, Inc.*, 1:13-CV-158, 2013 WL 6327852, *5–6 (E.D. Va. Dec. 3, 2013); *Dynetix Design Solutions, Inc. v. Synopsis, Inc.*, C 11-05973 PSG, 2013 WL 4538210, at *3–4 (N.D. Cal. Aug. 22, 2013); *AVM Techs., LLC v. Intel Corp.*, CIV.A. 10-610-RGA, 2013 WL 126233, at *3 (D. Del. Jan. 4, 2013). *Cf. Internet Machines LLC v. Alienware Corp.*, 6:10-CV-23, 2013 WL 4056282, at *13 (E.D. Tex. June 19, 2013) (“Because [the damages expert] used the smallest saleable unit as his royalty base, additional apportionment is unwarranted . . .”).

121 *Egry Register Co. v. Standard Register Co.*, 23 F.2d 438, 440 (6th Cir. 1928) (“[The patent owner] cannot, by the language which his claims happen to take, transform his invention of an improvement in an existing structure into one of a complete structure, as if it were wholly new, so as to entitle him to profits upon these parts of it which are not in any fair sense his invention.”).

real focus ought to be on the economic realities and not the vagaries of claim drafting,” particularly because “the way claims are drafted [is] . . . so manipulable.”¹²²

In determining how to properly apportion the royalty base, the focus should be on the point of novelty of the patented invention and not on how that invention is claimed.

Principle II-4: Where the accused product incorporates multiple technologies, once the proper royalty base has been determined, the reasonable royalty rate should reflect the relative contribution of the patented invention as compared to the other technologies incorporated into the royalty base. All technologies incorporated into the royalty base prior to trial should be considered. This approach should help to alleviate the problem often referred to as royalty stacking.

Comment

A product, especially a complex product, can, and often does, embody multiple different patents. Many electronic devices, including computers and tablets, as well as most software, and even certain pharmaceutical products, involve hundreds or even thousands of patents. For example, it is estimated that the 3G CDMA wireless standard involves nearly 1000 different patents and thus, 3G wireless handsets often involve large numbers of patents. The presence of royalty stacking, i.e., when multiple patents read on a single product, exacerbates the complexities of calculating a reasonable royalty. Ignoring the effects of royalty stacking may lead to an overestimation of actual patent damages and to overcompensation of the patent holder.¹²³ The Federal Circuit noted in *Integra Lifesciences I, Ltd. v. Merck KGaA*, that “[t]he cumulative effect of such stacking royalties can be substantial” and should play a role in determining patent damages.¹²⁴ Accordingly, with respect to a product in which multiple patents are potentially at issue, the reasonable royalty calculus should account for such royalty stacking.

A simple example helps illustrate the issue of royalty stacking. In a product covered by 25 different patents, assuming each were of equal value, if each patentee were awarded a 5% royalty, more than the entire revenue of the product is taken just to compensate for intellectual property rights.

From an end-results perspective, a royalty stacking-based calculation should arrive at the same conclusion as any other reasonable royalty calculation; namely, a stacking methodology seeks to reasonably attribute the contribution of the patent(s)-in-suit to the overall value of the product at issue. In short, patents covering a small component or feature of a larger invention should command a lower royalty rate than patents covering the whole product.

¹²² *FTC Report*, *supra* note 115, at 211.

¹²³ It should also be noted that while many commentators consider royalty stacking to be a significant concern, there are also several studies suggesting that royalty stacking may not be such a serious problem, or at least that its effect is not borne out in empirical studies. For instance, a 2008 article reviewed empirical literature from four separate industries – semiconductor, software, biomedical, and mobile phone – and concluded that royalty stacking was neither common nor costly enough to warrant policy changes at the litigation stage. Rather, the studies suggested that any potential harm was better handled through *ex ante* licensing strategies. See Damien Geradin et al., *The Complements Problem Within Standard Setting: The Evidence on Royalty Stacking*, 14 B.U. J. Sci. & Tech. L. 144, 145 (2008).

¹²⁴ *Integra Lifesciences I, Ltd. v. Merck KGaA*, 331 F.3d 860, 871–72 (Fed. Cir. 2003), *rev'd on other grounds*.

Recognizing the importance of stacking is particularly meaningful when assessing comparable licenses. In situations where royalty stacking is in play, analysis of comparable product licenses must recognize the contribution of the patent-at-issue to the overall contribution of value to the product-at-issue. Stated differently, in evaluating licenses, it is important not to confuse the notion of comparable licenses with respect to a given type of *product* with comparable licenses regarding a given *technology*. Thus, licenses that “stack” onto one product should be relevant and admissible even if not directly comparable.

Accounting for royalty stacking in assessing the reasonable royalty is easier said than done, however, given the amount and type of evidence that would be required at trial to elucidate the non-asserted patents that may read on the accused product(s). In a multi-featured product, courts have endorsed various ways of determining the value of the patented features, such as through the use of surveys, conjoint analysis, hedonic regression, or other analytical methods that suit the circumstances. Such evidence relating to non-asserted patents will complicate the trial but is necessary to arrive at a reasonable royalty. The party arguing that stacking is a concern should come forward with other licenses or royalty demands on the product-in-suit. Allegations of stacking without such evidence would require proving other patents read on the product at issue where no license was sought, and is too speculative. Such proof would result in a patent trial within a patent trial and would quickly become unmanageable.

Principle II-5: A reasonable royalty must reflect the extent to which, throughout the period of infringement, the patented invention has represented an improvement over available alternatives at the time of infringement, including the prior art. A royalty which over- or under-values the inventive contribution of the patent claim is not reasonable.

Comment

The reasonable royalty should compensate the patent owner for the claimed invention’s incremental improvement over the prior art or the next best commercially acceptable non-infringing alternative, if such alternative exists. This notion is consistent with economic theory in that the value of an asset is limited to the economic benefit that can be obtained from using the asset vis-à-vis that which can be obtained by using the next best available alternative. It is the incremental economic benefit over that which can be earned using the next best alternative that determines the economic value of the patented invention, and thus the royalty the owner can obtain in the market for the use of that patented invention.

The principle that damages due a patent owner should be commensurate with the incremental economic contribution afforded by the claimed invention(s) was recognized in *Grain Processing Corp. v. American Maize Products*, in which the Federal Circuit held that “only by comparing the patented invention to its next-best available alternative(s) – regardless of whether the alternative(s) were actually produced and sold during the infringement – can the court discern the market value of the patent owner’s exclusive right, and therefore his expected profit or reward, had the infringer’s activities not prevented him from taking full economic advantage of this right.”¹²⁵

125 *Grain Processing Corp. v. American Maize Products Co.*, 185 F.3d 1341, 1351 (Fed. Cir. 1999).

A conceptual problem may arise when the claimed invention is comprised of a combination of prior art elements. However, in such instances, it is the incremental functionality enabled by the unique combination of prior art elements that comprises the novelty, which is then compared to the prior art for purposes of determining a reasonable royalty.

Even where the claimed invention is but one of multiple components or features within a larger apparatus and is shown to be the basis, or substantially all of the basis, of consumer demand for the entire apparatus as per the entire market value rule, the incremental value of what is contributed by the invention must be assessed. In such instances, the reasonable royalty appropriately should reflect the economic value associated with the entire market value of the apparatus, measured by the benefit of the apparatus containing the invention over the next best alternative to the apparatus.

Principle II-6: Three principles apply to the consideration of a non-infringing alternative design, or “design around,” in determining a reasonable royalty, as noted below:

Principle II-6(a): Evidence of a non-infringing design around that is technically and commercially feasible and available during the damages period is relevant to the reasonable royalty determination. A design around need not actually have been implemented in order to be considered, but must be raised during fact discovery to prevent expense, delay, and prejudice.

Principle II-6(b): In order to be considered a design around in the first instance, the proposed alternative design must be shown by a preponderance of the evidence not to infringe the asserted claims of the patent(s)-in-suit, and to be an acceptable substitute.

Principle II-6(c): On a proper showing, the total economic cost of the infringer’s next best available alternative may serve to cap the damages award.

Comment

Courts have long considered the availability of non-infringing alternatives to be well within the scope of information relevant to the determination of a reasonable royalty.¹²⁶ Indeed, it is often analyzed as a *Georgia-Pacific* factor.¹²⁷ Thus, under existing law, it is clear that a legitimate non-infringing design around alternative may be considered in the reasonable royalty analysis.

¹²⁶ See, e.g., *Panduit Corp. v. Stahlbr Bros. Fibre Works, Inc.*, 575 F.2d 1152, 1159 (6th Cir. 1978); *Hughes Tool Co. v. G.W. Murphy Industries, Inc.*, 491 F.2d 923, 931 (5th Cir. 1973); *Smithkline Diagnostics, Inc. v. Helena Labs. Corp.*, No. B-83-10-CA, 1989 WL 418791, at *6 (E.D. Tex. June 30, 1989), *aff’d* 926 F.2d 1161 (Fed. Cir. 1991); see also *Radio Steel & Mfg. Co. v. MTD Prods., Inc.*, 788 F.2d 1554, 1557 (Fed. Cir. 1986); *Columbia Wire Co. v. Kokomo Steel & Wire Co.*, 194 F.108, 110 (7th Cir. 1911).

¹²⁷ *Georgia-Pacific*, 318 F. Supp. at 1120 (listing “the utility and advantages of the patent property over the old modes or devices, if any, that had been used for working out similar results,” as the ninth of fifteen evidentiary factors for determining a reasonable royalty); see also *i4i Ltd. P’ship v. Microsoft Corp.*, 598 F.3d 831 (Fed. Cir. 2010) (affirming admissibility of expert’s reliance on *Georgia-Pacific* factors in his reasonable royalty determination, including the lack of “acceptable non-infringing alternatives . . . at the time of the hypothetical negotiation”); *LaserDynamics, Inc. v. Quanta Computer, Inc.*, No. 2:06-cv-348, 2011 WL 197869, at *2 (E.D. Tex. Jan. 20, 2011); *Abbott Labs. v. Sandoz, Inc.*, 743 F. Supp. 2d 762, 773 (N.D. Ill. 2010).

It is often the case, however, that conclusory assertions of available design arounds are made after the close of fact discovery through a damages expert during the expert discovery period. Such late and general assertions lead to *Daubert* motions, motions *in limine*, and/or evidentiary objections at trial that require the court to make an important evidentiary determination late in the proceedings, and with little information. Further, if the court permits the introduction of evidence of the late-identified design around, the patent holder may be at a distinct disadvantage. The patent holder may not be in a position to mount a persuasive challenge as to the technical or commercial merits of the asserted design around due to the lack of fact discovery, yet the argument that the asserted design around should limit the damages award may have appeal to the fact-finder. Not only is there uncertainty for both litigants and the court regarding the required timing and extent of the disclosure of any potential design around, the evidentiary standards regarding the admissibility of any asserted design around are also unclear.

In the context of lost profits, the initial burden is on the patent holder to show, *inter alia*, the absence of acceptable non-infringing alternatives. Once the absence of acceptable non-infringing alternatives is proven, the burden of proof shifts to the accused infringer, who must then prove that an acceptable non-infringing alternative exists.¹²⁸ While the law as to who bears the initial and ultimate burden in the reasonable royalty context is not settled,¹²⁹ there is no reason for it to differ. That is, the patent holder should initially evaluate any non-infringing alternatives proffered by the accused infringer during discovery, and show why they are infringing and/or not acceptable substitutes. The burden should then shift to the accused infringer to prove that the proffered alternative is, in fact, non-infringing and acceptable so as to mitigate the potential damages award.

Common sense and economics both argue that, in general, it would be irrational for an accused infringer to pay more for a license to a patent than the total economic cost it would incur to implement its next best available alternative to the patented technology, inclusive of all of the costs associated with the implementation of that next best alternative (such costs would necessarily include any costs due to any inferiorities of the alternative as compared to the patented technology). In the real world, negotiators for patent licenses consider the alternatives to patent licenses, including achieving the goal of the patented technology in a manner that does not require a license.¹³⁰ So, too, goes the argument, should the hypothetical negotiators.¹³¹

-
- 128 See, e.g., *Zygo Corp. v. Wyko Corp.*, 79 F.3d 1563, 1571 (Fed. Cir. 1996) (noting that the plaintiff bears the burden on lost profit damages calculations, including the non-existence of a non-infringing alternative); *Smithkline Diagnostics, Inc. v. Helena Labs. Corp.*, 926 F.2d 1161, 1165 (Fed. Cir. 1991); *TWM Mfg. Co. v. Dura Corp.*, 789 F.2d 895, 901 (Fed. Cir. 1986); see also *Grain Processing*, 185 F.3d at 1354 (finding that a non-infringing alternative need not be on the market during the infringement period to factor into a lost profits analysis, but nonetheless, “[w]hen an alleged alternative is not on the market during the accounting period, a trial court may reasonably infer that it was not available as a non-infringing substitute at that time,” which means then that the burden then falls on the infringer to prove availability, and the fact-finder “must proceed with caution” in assessing that proof).
- 129 See, e.g., *ResQNet.com*, 594 F.3d at 872 (“The district court seems to have been heavily influenced by Lansa’s decision to offer no expert testimony to counter Dr. David’s opinion. But it was ResQNet’s burden, not Lansa’s, to persuade the court with legally sufficient evidence regarding an appropriate reasonable royalty. As a matter of simple procedure, Lansa had no obligation to rebut until ResQNet met its burden with reliable and sufficient evidence. This court should not sustain a royalty award based on inapposite licenses simply because Lansa did not proffer an expert to rebut Dr. David.” (citing *Lucent*, 580 F.3d at 1329 (“Lucent had the burden to prove that the licenses were sufficiently comparable to support the lump sum damages award.”)); THE NAT’L JURY INSTRUCTION PROJECT, MODEL PATENT JURY INSTRUCTIONS 64–66 (2009), available at <http://www.nationaljuryinstructions.org/documents/NationalPatentJuryInstructions.pdf> (providing clear burden of proof on plaintiff for establishing lost profits damages, but listing a number of considerations for the jury in determining reasonable royalty damages, where the plaintiff did not establish lost profit damages for any or all of the accused products).
- 130 See, e.g., C.W. Shifley, Commentary, *Alternatives to Patent Licenses: Real-World Considerations of Potential Licenses Are -- and Should Be -- a Part of the Courts’ Determinations of Reasonable Royalty Patent Damages*, 34 IDEA: J.L. & TECH. 1, 3 (1993). Mr. Shifley quotes George E. Frost, former patent counsel for General Motors, describing the real world process of negotiating a license royalty as necessitating a determination of the “increment of value”: “[W]e need to be sure that we don’t get into royalties that are more than what the alternative costs, because there’s no sense at all to pay more to use the patent including the royalty, . . . than it would cost to use the alternative.” *Id.*
- 131 See, e.g., *id.* at 2 (quoting commentator George E. Frost as saying “[t]he dollar disadvantage of going to the most practical non-infringing alternative in lieu of the patent product . . . places a ceiling on what any rational negotiator would . . . pay”).

This economic theory has become established in the consideration of a lost profits award. In *Grain Processing Corp. v. American Maize Products Co.*, the Federal Circuit held that the existence of a non-infringing alternative must be considered in reconstructing the “but for” world, and thus may cap or render unavailable a lost profits damages award.¹³² Although the Federal Circuit seemed poised to apply *Grain Processing* in the reasonable royalty context,¹³³ it declined the opportunity to do so in *Mars, Inc. v. Coin Acceptors, Inc.* In *Mars*, the accused infringer argued that “an infringer should not be required to pay more in reasonable royalties than it would have paid to avoid infringement in the first place by switching to an available non-infringing alternative.”¹³⁴ The Federal Circuit rejected this argument, both because “there was no available and acceptable non-infringing alternative to which Coinco could have switched at the time of the hypothetical negotiation,” and because “it is wrong as a matter of law to claim that reasonable royalty damages are capped at the cost of implementing the cheapest available, acceptable, non-infringing alternative.”¹³⁵ Instead, the Federal Circuit stated, “[t]o the contrary, an infringer may be liable for damages, including reasonable royalty damages that exceed the amount that the infringer could have paid to avoid infringement.”¹³⁶ Although the *Mars* court rejected the notion that reasonable royalty damages are *always* capped by the cost of implementing a non-infringing alternative, it left open the possibility that reasonable royalty damages may be capped by such a cost in the appropriate circumstances.

In keeping with *Mars* and in an effort to move the law forward, the Working Group favors limiting the reasonable royalty award, in certain specific circumstances and upon a proper showing, to the aggregate cost associated with implementing a technically and commercially feasible non-infringing alternative in place of the accused instrumentality. Such a limitation is supported by economic theory,¹³⁷ and is consistent with the retrospective paradigm of the hypothetical negotiation articulated above.

The total economic cost to implement an available non-infringing alternative would include, for example, R&D expense, product development expense, any incremental manufacturing costs, and any foregone profits due to time-to-market considerations as well as changes in prices and market share.

-
- 132 See *Grain Processing*, 185 F.3d at 1350–51 (“[A] fair and accurate reconstruction of the ‘but for’ market also must take into account, where relevant, alternative actions the infringer foreseeably would have undertaken had he not infringed. Without the infringing product, a rational would-be infringer is likely to offer an acceptable non-infringing alternative, if available, to compete with the patent owner rather than leave the market altogether. The competitor in the ‘but for’ marketplace is hardly likely to surrender its complete market share when faced with a patent, if it can compete in some other lawful manner.”).
- 133 See *Riles v. Shell Exploration & Prod. Co.*, 298 F.3d 1302, 1312 (Fed. Cir. 2002) (“The economic relationship between the patented method and non-infringing alternative methods, of necessity would limit the hypothetical negotiation.”) (citing *Grain Processing*, 185 F.3d at 1347 (the difference in production costs between infringing and non-infringing products “effectively capped the reasonable royalty award”)); *Zygo*, 79 F.3d at 1571–72 (Fed. Cir. 1996) (vacating and remanding the district court’s damages determination where both lost profits and reasonable royalty failed to take into account non-infringing alternatives that would have given the alleged infringer a “stronger position to negotiate for a lower royalty rate”); see also *Micro Chemical, Inc. v. Lextron, Inc.*, 317 F.3d 1387, 1393 (Fed. Cir. 2003) (reserving judgment on “whether the holding of *Grain Processing* has applicability in the reasonable royalty context”).
- 134 *Mars, Inc. v. Coin Acceptors, Inc.*, 527 F.3d 1359, 1373 (Fed. Cir. 2008).
- 135 *Id.* Interestingly, the Federal Circuit did not cite or mention *Grain Processing* in the opinion.
- 136 *Id.*; see also *Boeing Co. v. United States*, 86 Fed. Cl. 303, 319 n.14 (2009) (“The Federal Circuit recently rejected the argument that ‘reasonabl[e] royalty damages are capped at the cost of implementing the cheapest available, acceptable, non-infringing alternative.’”).
- 137 See, e.g., Joan L. Eads, Commentary, *Does Grain Processing Apply in a Reasonable Royalty Damage Analysis?*, 10 No. 26 ANDREWS INTEL. PROP. REP. 13 (2004); ECONOMIC APPROACHES TO INTELLECTUAL PROPERTY 52–57 (Gregory K. Leonard & Lauren J. Stiroh eds., 2005); Nathaniel C. Love, Comment, *Nominal Reasonable Royalties for Patent Infringement*, 75 U. CHI. L. REV. 1749, 1757–66 (2008); Liane M. Peterson, Grain Processing and Crystal Semiconductor: *Use of Economic Methods in Damage Calculations Will Accurately Compensate For Patent Infringement*, 13 FED. CIRCUIT B.J. 41 (2003); John W. Schlicher, *Measuring Patent Damages by the Market Value of Inventions – The Grain Processing, Rite-Hite, and Aro Rules*, 82 J. PAT. & TRADEMARK OFF. SOC’Y 503 (2000); Christopher B. Seaman, *Reconsidering the Georgia-Pacific Standard for Reasonable Royalty Patent Damages*, 2010 BYU L. REV. 1661, 1711–26; John S. Torkelson, *Calculating Reasonable Royalty Damages for Infringement of Early-Stage Technology Patents*, 4 SEDONA CONF. J. 47, 56–57 (2003); see also *FTC Report*, *supra* note 115, at 160–76 (noting that where courts “reject, either implicitly or explicitly, a limitation based on the maximum amount a willing licensee would pay,” they “risk overcompensating patentees in litigation as compared to the market and creat[e] problems such as higher prices, increased patent speculation, and decreased innovation”).

However, to avoid undercompensating the patent holder, the court should require the litigants to follow certain best practices outlined below before allowing the accused infringer to present an argument that an asserted non-infringing design around alternative should cap the reasonable royalty award.

BEST PRACTICES

1. *The fact-finder should make the ultimate determination as to whether a proffered design around alternative is both non-infringing and acceptable by evaluating the evidence with the understanding that it is the accused infringer who bears the ultimate burden of establishing an acceptable design around by a preponderance of the evidence.*
2. *The proponent of a royalty cap based on the existence of a non-infringing design around alternative must provide competent, admissible evidence (subject to cross-examination) regarding:*
 - a. *All costs associated with implementation of the design around over the implementation of the accused instrumentality, measured at the time the accused instrumentality was in the planning stages, in other words, the additional amount it would have cost the accused infringer to have implemented the non-infringing design rather than the accused instrumentality, e.g., design and development costs (including any additional personnel costs), cost of materials, and costs associated with any required redesign of other components to accommodate the non-infringing design (including any licensing fees for implementing a third party's patented technology in the non-infringing design).*
 - b. *All costs associated with the marketing and selling of the design around over the marketing and selling of the accused instrumentality, e.g., sales force training that would be required for the design around that was not required for the accused instrumentality.*
 - c. *The technical and commercial equivalence of the design around to consumers such that consumer demand for the design around would have been equivalent to the consumer demand for the accused instrumentality.*

Principle II-7: **Where the technology claimed in the asserted patent is necessary to practice because (1) it is essential to a *de facto* standard or a standard adopted by a recognized standard setting organization (i.e., standard-essential); (2) a technically feasible non-infringing design around alternative is restricted or prohibited by government regulations or requirements; and/or (3) the technically feasible design around is cost-prohibitive, then the reasonable royalty should exclude any premium the patent may command solely resulting from the adoption of the standard or the governmental/commercial prohibitions on design modification. All standards adopted prior to trial may be considered.**

Comment

As recognized by the court in *Microsoft Corp. v. Motorola, Inc.*, “when [a] standard becomes widely used, the holders of Standard Essential Patents (SEPs) obtain substantial leverage to demand more than the value of their specific patented technology. This is so even if there were equally good alternatives to that technology available when the original standard was adopted. After the standard is widely implemented, switching to those alternatives is either no longer viable or would be very costly.”¹³⁸ Many commentators have described patent holdup – i.e., the ability of SEP owners to demand more than the value of their patented technology, and to attempt to capture the value of the standard itself – as a serious factor in increased licensing rates or royalty calculations, which can result in overcompensation of patent holders.¹³⁹

Potential holdup or lock-in effects arise primarily in two scenarios. In the first, a patented component has become necessary to practice a technical standard. The technical standard may be one that was promulgated by a standard setting organization or it may simply be a standard that has arisen as the result of market forces (commonly referred to as a *de facto* standard).¹⁴⁰ The second holdup or lock-in scenario is presented when a technically feasible, non-infringing alternative exists for a component of a product, but that alternative cannot practically be incorporated into the product because the cost of redesign would be commercially prohibitive.

As indicated in the discussion above, when the Working Group selected the Retrospective Model for the application of patent damages, we recognized that a departure from this model would be required to address the problem of lock-in. That is, for patent infringement that does not involve lock-in and the concomitant potential for hold-up, the Retrospective Model places valuation of the patented technology at the time of trial. However, where lock-in effects exist at the time of trial, the valuation of the patented technology must be performed at an earlier time, before the infringer was locked-in, so as to avoid the attachment of a premium to the value of the patent technology that results from the user’s lock-in. Accordingly, the Working Group determined that, for purposes of addressing lock-in and avoiding holdup effects, the patented technology to which the infringer is locked in generally should be valued in a manner that would exclude any premium the patent would command as a result of the adoption of the standard, i.e., any premium divorced from the technical merits of the technology.

This approach is consistent with that articulated in the FTC Report: “A reasonable royalty damages award that is based on high switching costs, rather than the ante value of the patented technology compared to alternatives, overcompensates the patentee To prevent damage awards based on switching costs, courts should set the hypothetical negotiation at an early stage of product development, when the infringer is

138 *Microsoft Corp. v. Motorola, Inc.*, No. C10-1823JLR, 2013 WL 2111217, at *10 (W.D. Wash., Apr. 25, 2013); see also *In re Innovatio IP Ventures*, 2013 WL 5593609, at *8–9, *40–41.

139 See, e.g., Thomas F. Cotter, *Patent Holdup, Patent Remedies, and Antitrust Responses*, 34 J. CORP. L. 1151 (2009); Joseph Scott Miller, *Standard Setting, Patents, and Access Lock-In: RAND Licensing and the Theory of the Firm*, 40 IND. L. REV. 351 (2007); Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007); Carl Shapiro, *Injunctions, Hold-Up, and Patent Royalties*, 12 AM. L. & ECON. REV. 280 (2010); but see Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMP. L. & ECON. 535 (2008); J. Gregory Sidak, *Holdup, Royalty Stacking, and the Presumption of Injunctive Relief for Patent Infringement: A Reply to Lemley and Shapiro*, 92 MINN. L. REV. 714 (2008).

140 Examples of such *de facto* standards include the QWERTY keyboard and the MP3 audio format. For more regarding the development and effects of *de facto* standards, see Janice M. Mueller, *Patent Misuse Through the Capture of Industry Standards*, 17 BERKELEY TECH. L.J. 623, 633–35 (2002).

making design decisions.”¹⁴¹ Similarly, the FTC Report advocates for valuation of standards-essential patents “based on the ex ante value of the patented technology at the time the standard is set.”¹⁴²

In Chapter III, the Working Group recommends the best practices when calculating reasonable royalty patent damages in cases presenting the above-mentioned scenarios.¹⁴³

Principle II-8: The comparison of any proposed comparable license to the hypothetical license should itemize and separately value – to the extent possible – the material ways in which the two differ.

Comment

The first *Georgia-Pacific* factor considers “[t]he royalties received by the patentee for the licensing of the patent in suit, proving or tending to prove an established royalty.”¹⁴⁴ The second *Georgia-Pacific* factor is “the rates paid by the licensee for the use of other patents comparable to the patent in suit.”¹⁴⁵ While *Georgia-Pacific* factors 1, 2 and 12 allow the expert witness to consider certain license agreements in determining the proper royalty rate and royalty structure, courts have not provided a definitive, comprehensive outline stating what criteria must be evaluated to determine if a license agreement is properly “comparable.”

The Federal Circuit addressed comparability of license agreements in *Lucent*, during its consideration of *Georgia-Pacific* factor 2.¹⁴⁶ The Federal Circuit analyzed various license agreements relied upon by Lucent in its presentation of royalty damages figures to the jury, and found that the evidence presented at trial was not sufficient for Lucent to have met its burden that the licenses were sufficiently comparable to support the damages award.¹⁴⁷ The Federal Circuit considered multiple factors regarding the licenses’ comparability to the hypothetical negotiation for the patents-in-suit, including the similarity and importance of the technology licensed, the price of the licensed product, the complexity of the royalty rate, and the structure of payment (i.e., lump sum royalty payments as opposed to running royalty payments).¹⁴⁸

Although the *Lucent* decision does not forbid the use of license agreements that differ, even substantially, from the hypothetical license, including differences in the technology covered in the licenses versus the patents-in-suit, or differences in the royalty structure (e.g., lump sum versus running royalty), it does highlight the importance of having the damages expert witness conduct a rigorous comparison of the various terms of the real-world licenses to the hypothetical licenses so as to justify use of dissimilar real-life license agreements.¹⁴⁹

141 *FTC Report*, *supra* note 115, at 190–91.

142 *Id.* at 194. The FTC Report goes on to recommend “Courts should cap the royalty [for patents subject to a RAND commitment] at the incremental value of the patented technology over other alternatives available at the time the standard was defined.” *Id.* See also *In re Innovatio IP Ventures*, 2013 WL 5593609, at *40–41 (“The court must, however, not consider the effect of standardization when evaluating the *ex ante* negotiation in 1997.”).

143 See *infra*, Chapter III, Best Practice Nos. 9, 15, and accompanying commentary.

144 *Georgia Pacific Corp. v. U.S. Plywood Corp.*, 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970).

145 *Id.*

146 *Lucent*, 580 F.3d at 1325–32.

147 *Id.* at 1332.

148 *Id.* at 1327–31 (addressing each of eight real-life licenses presented by Lucent’s damages expert to the jury).

149 See, e.g., *id.* at 1330 (“As we noted above, certain fundamental differences exist between lump sum agreements and running royalty agreements. This is not to say that a running royalty license agreement cannot be relevant to a lump sum damages award, and vice versa. For a jury to use a running royalty agreement as a basis to award lump sum damages, however, some basis for comparison must exist in the evidence presented to the jury. In the present case, the jury had almost no testimony with which to recalculate in a meaningful way the value of any of the running royalty agreements to arrive at the lump sum damages award.”).

The Federal Circuit also addressed comparability of license agreements in *ResQNet.com, Inc. v. Lansa, Inc.*¹⁵⁰ and *Wordtech Sys., Inc. v. Integrated Network Solutions, Inc.*¹⁵¹ These decisions, also highlight the importance of evaluating and comparing the circumstances and considerations of a potentially comparable license agreement to the facts and circumstances of the hypothetical negotiation for the patent(s)-in-suit, including, but not limited to, demonstrating a link between the invention involved in a potentially comparable license agreement to the claimed invention in the patent(s)-in-suit. Furthermore, in *ResQNet.com*, the Federal Circuit criticized an expert's reliance on license agreements that included payments for add-ons, such as marketing and other services, unrelated to the licensed technology.¹⁵²

District court decisions vary in their approaches. Some articulate a standard similar to that of the Federal Circuit in *ResQNet*.¹⁵³ Other district courts have been less strict in analyzing what constitutes a comparable license.¹⁵⁴

The Working Group supports a rigorous analysis to determine the comparability of license agreements. Rigorously analyzing and adjusting for any material differences between a benchmark license and the hypothetical license provides a rational and justifiable basis for determining what royalty would result from the hypothetical negotiation.

This proposed construct of analyzing and adjusting for differences between a benchmark asset and the asset being valued is common practice in valuation analyses, and is analogous to adjusting the estimated value of a parcel of real estate based on the differing characteristics of a comparable, recently-sold parcel. To the extent that both properties are identical with respect to a given characteristic (e.g., square footage), no adjustment to the estimated value is necessary. Conversely, to the extent the properties differ with respect to a given characteristic (e.g., more desirable location), an adjustment to the estimated value may be warranted.

The factors one might consider in assessing the comparability, and therefore the probative value, of a benchmark license to the hypothetical license will vary, given the unique facts and circumstances of each license. However, the comparability factors that one would likely consider will fall within four primary categories: (1) comparability of the licensed technologies; (2) comparability of the terms of the licenses; (3) comparability of the commercial and legal circumstances in which the parties negotiated the licenses; and (4) the bona fide nature of the license, including whether royalties were actually paid under the agreement, and the circumstances surrounding any non-payment. In those instances where comparability factors differ as between a proposed comparable license and the hypothetical license, the reasonable royalty analysis should attempt to quantify as accurately as possible the amount by which the indicated royalty should be adjusted. In those instances where it is not possible to quantify the requisite adjustment, the reasonable royalty analysis should, nonetheless, indicate and reflect the general nature that each comparability factor would have on the indicated royalty (e.g., whether it would tend to raise or lower the royalty).

In Principle 8(a), we address bare patent licenses to the patent(s)-in-suit in the same field of use, i.e., a one-way license to patents with no cross-license or other technology

150 594 F.3d 860 (Fed. Cir. 2010).

151 609 F.3d 1308, 1320 (Fed. Cir. 2010).

152 *ResQNet.com*, 594 F.3d at 870.

153 See, e.g., *IP Innovation*, 705 F. Supp. 2d at 691; *Fenner Invs.*, 2010 U.S. Dist. LEXIS 101384, at *5; *Ricoh*, 2010 U.S. Dist. LEXIS 27301 at *26–27.

154 See, e.g., *Saffran*, 2011 U.S. Dist. LEXIS 34858, at *31–32; *Finjan Software*, 2009 U.S. Dist. LEXIS 72825, at *41.

transfer. Bare patent licenses to the patent(s)-in-suit in the same field of use may provide guidance to the appropriate royalty rate determination from a hypothetical negotiation, as the technology is identical; however, as addressed below, a proper and rigorous analysis is required to determine if the agreement is truly comparable.

In Principle 8(b), we address licenses that are not a bare patent license to the patent(s)-in-suit in the same field of use, but which license the patent(s)-in-suit as part of a more comprehensive license.

In Principle 8(c), we address licenses that do not involve the patent(s)-in-suit, but which may nevertheless provide guidance to the appropriate royalty rate reached in a hypothetical negotiation, if subjected to a rigorous analysis.

In Principle 8(d), we address license agreements arising from litigation settlements.

With respect to all scenarios encompassed by principles 8(a) – 8(d), any prior license agreement that lacks sufficient indicia of comparability to the hypothetical license should be disregarded in a reasonable royalty analysis.

Principle II-8(a): Evidence of a bare patent license to the patent(s)-in-suit in the same field of use as the accused product/service is generally relevant to the reasonable royalty inquiry and should usually be considered in the determination of the reasonable royalty.

Comment

To be considered in the Retrospective Model, the license may have been entered into any time prior to trial.

When a bare patent license to the patent(s)-in-suit in the same field of use as the accused product/service exists, certain of the market factors, license terms and technological considerations listed in the section below may be relevant to the determination of the reasonable royalty. The ones relevant to this principle are so indicated.

Principle II-8(b): On a proper showing, evidence of a license that is not a bare patent license to the patent(s)-in-suit in the same field of use, but which does license the patent(s)-in-suit as part of a more comprehensive license, may be relevant to the reasonable royalty inquiry.

Comment

To be considered in the Retrospective Model, the license may have been entered into any time prior to trial.

When a license that is not a bare patent license to the patent(s)-in-suit in the same field of use as the accused product/service but rather is a license to the patent(s)-in-suit as part of a more comprehensive license, certain of the below listed license terms, market

factors and technology considerations may be relevant to the determination of the reasonable royalty. The ones relevant to this principle are so indicated.

Principle II-8(c): On a proper showing, evidence of a license that does not license the patent(s)-in-suit may be relevant to the reasonable royalty inquiry.

Comment

To be considered in the Retrospective Model, the license may have been entered into any time prior to trial.

A license agreement that does not license the patent(s)-in-suit may (or may not) be relevant to the determination of the reasonable royalty rate depending on the outcome of a rigorous analysis that compares and contrasts the proposed license agreement to the hypothetical license, including an analysis of certain of the below listed technology considerations, the relevant product/service, license terms and market factors. The ones relevant to this principle are so indicated.

Principle II-8(d): On a proper showing, license agreements in settlement of litigation that license the patent(s) or technology-in-suit may be relevant to the reasonable royalty determination.

Comment

Courts have considered both the admissibility and the use of settlement agreements as comparables in patent damages determinations. Some district courts have denied the use of settlement agreements due to potential jury confusion and prejudice.¹⁵⁵ Other district courts have allowed the use of settlement agreements on a case-by-case basis if the settlement agreements are for the patent(s)-in-suit, are the only sufficiently comparable license agreement(s), and resemble agreements negotiated outside of litigation.¹⁵⁶ The Working Group supports a rigorous analysis to determine the comparability of settlement agreements and their use at trial.

To be considered in the Retrospective Model, the settlement license may have been entered into any time prior to trial.

A settlement license that licenses the patent(s)-in-suit may be relevant to the determination of the reasonable royalty rate depending on the outcome of a rigorous analysis that compares and contrasts the proposed settlement agreement to the hypothetical license including an analysis of certain of the below listed potential factors related to settlements of litigation, relevant license terms, product/service, market factors and technology considerations. The ones relevant to this Principle are so indicated.

¹⁵⁵ *Fenner Inv.*, 2010 U.S. Dist. LEXIS 41514, at *10; *Cornell Univ.*, 2008 U.S. Dist. LEXIS 39343, at *13–14.
¹⁵⁶ *ResQNet.com*, 594 F.3d at 872; *ReedHycalog*, 727 F. Supp. 2d at 546 (E.D. Tex. 2010).

FACTORS AND CONSIDERATIONS RELEVANT TO COMPARABILITY ANALYSIS

1. *Cross-license provisions*: Unless the potentially comparable license distinguishes/breaks-out the consideration/royalty of the intellectual property of each entity separately, licenses that include cross-license provisions are generally not useful for determination of the reasonable royalty rate. However, depending on the facts and circumstances, licenses that include cross-license provisions may still provide guidance for other damage related factors, including the appropriate royalty base and royalty structure (e.g., per unit v. running royalty). (Relevant to Principles II-8(b), 8(c), and 8(d)).
2. *Additional patents to patent(s)-in-suit*: If the potentially comparable license licenses the patent(s)-in-suit plus additional patents, but does not distinguish/break-out the consideration/royalty for the patent(s)-in-suit from the additional patents in the license, and the record does not provide guidance as to the licensor's and licensee's perceptions of the value of the patent(s)-in-suit separate from the additional patents, it may not be possible to determine the portion of the consideration/royalty attributable to the patent(s)-in-suit. However, depending on the facts and circumstances of the case, the consideration/royalty in the potentially comparable license may tend to indicate an upper bound to the royalty rate determination in the hypothetical negotiation, as the licensee in the hypothetical negotiation would not be granted rights to the additional patents. Depending on the facts and circumstances, the potentially comparable license may also provide guidance for other damage related factors, including the appropriate royalty base and royalty structure. (Relevant to Principles II-8(b) and 8(d)).
3. *Different Patents/Technology from patent(s)-in-suit*: Different patents in the same technology "field" may have very different values, even where they are used in the same products. Thus, if the potentially comparable license contains a license to patents that are different from the patents(s)-in-suit, it may not be possible to reliably use the license as a comparable license. To assess comparability, as well as to quantify the relative royalty rates (after a determination that it can be done reliably), it is important to consider: (1) the relative importance of the licensed technology to the accused product/service in the potentially comparable license as compared to the relevant product/service in the hypothetical negotiation, including the importance of the licensed technology to demand, sales, profits and price of the relevant product/service; (2) whether the licensed technology is related to only one component of the overall relevant product/service and the number; and (3) importance of other technologies included in the relevant products/services. (Relevant to Principles II-8(b), 8(c), and 8(d)).
4. *Additional licensed property*: (e.g., trade secrets, know-how, technical assistance). If the potentially comparable license does not distinguish/break-out the consideration/royalty for each licensed

property, and the record does not provide guidance as to the licensor's and the licensee's perceptions of the value of each licensed property, it may not be possible to determine the portion of the consideration/royalty attributable to the patent(s)-in-suit. However, depending on the facts and circumstances of the case, the consideration/royalty in the potentially comparable license may tend to indicate an upper bound to the royalty rate determination in the hypothetical negotiation, as the licensee in the hypothetical negotiation would not be granted the additional rights. Depending on the facts and circumstances, the potentially comparable license may also provide guidance for other damage related factors including the appropriate royalty base and royalty structure. (Relevant to Principles II-8(b), 8(c), and 8(d)).

5. *Additional business arrangements associated with the potentially comparable agreement:* (e.g., development agreements, marketing agreements, and supply agreements). Even if the potentially comparable license distinguishes/breaks-out the consideration/royalty of the patent(s)-in-suit separately from additional business arrangements between the licensee and licensor, licenses that include or are related to other agreements that include additional business arrangements may not be useful for determination of the reasonable royalty rate, as it may not be possible to determine what royalty rate would have been negotiated but for the additional business arrangements. Additionally, if the licensor and licensee are related parties, or one party is a supplier of the other party, the royalty rate in the potentially comparable license may not be an "arms-length" transaction. However, depending on the facts and circumstances, license (or settlement) agreements that include additional business arrangements may still provide guidance for other damage related factors, including the appropriate royalty base and royalty structure. (Relevant to Principles II-8(b), 8(c), and 8(d)).
6. *Relevant product/service:* When the potentially comparable license is related to a different product/service than what is accused in the hypothetical negotiation, it is important to consider and compare the sales, profits, and price of the relevant products/services over the relevant period, the industry and competitive market for the relevant product/services, technological considerations (see below), and the relative importance of the relevant product/service to the overall business of the licensee. (Relevant to Principles II-8(b), 8(c), and 8(d)).
7. *Bargaining position of the parties/economic considerations:* It is appropriate to consider the relative bargaining positions of the licensor and licensee of the potentially comparable license to the bargaining positions of the licensor and licensee in the hypothetical negotiation. Potential relevant factors to consider include market position (e.g., competitors), anticipated sales volumes, anticipated profitability, anticipated market share, importance of the accused product/service to the overall business of the licensee, the business relationships of the parties (e.g., if one party is a supplier of the

accused product), and the extent to which royalties were paid/obtained under the potentially comparable license (or settlement). (Relevant to Principles II-8(a), 8(b), 8(c), and 8(d)).

8. *Royalty structure*: It may be appropriate to consider the licensing practices of the patent holder, the alleged infringer, and the industry regarding royalty structure to the extent it is different from the royalty structure of the potentially comparable license (or settlement). Depending on the facts and circumstances, it may be appropriate to convert a lump sum royalty in a potentially comparable license to a running royalty for purposes of application to the hypothetical license, if there is evidence of the parties' expected volume of relevant sales. Similarly, it may be appropriate to convert a running royalty into a lump sum royalty where the facts and circumstances suggest that the accused infringer would have agreed to a lump sum royalty. (Relevant to Principles II-8(a), 8(b), 8(c), and 8(d)).
9. *Date of the agreement and term*: It may be appropriate to consider changes in the relevant market of the accused product/service over time if the potentially comparable license (or settlement) was negotiated at a different point in time or for a different length of time than the hypothetical negotiation, including, for example, factors such as price, sales and profitability of the accused product/service, competition, industry standards, regulatory changes, and adoption of new technology into the accused product/service. (Relevant to Principles II-8(a), 8(b), 8(c), and 8(d)).
10. *Scope of the license*: It may be appropriate to consider the scope of the license of the potentially comparable license to the scope of the license in the hypothetical negotiation including, for example, consideration of territory (U.S. v. worldwide), exclusivity, rights granted (e.g., make, use, sell, and sublicense), and field of use. (Relevant to Principles II- 8(a), 8(b), 8(c), and 8(d)).
11. *Assumption (or not) of Validity and Infringement*: It may be appropriate to consider whether the potentially comparable license inherently or explicitly includes an assumption that the licensed patents are valid and infringed. The hypothetical license assumes validity and infringement, whereas license agreements and settlement agreements may be premised on a more uncertain picture of these merits-based issues. (Relevant to Principle II-8(d)).¹⁵⁷
12. *Litigation Factors*: In addition to consideration of the foregoing factors when evaluating a litigation settlement for comparability, it is appropriate to consider the facts and circumstances relating to the

¹⁵⁷ However, statements in litigation settlements (or other licenses, for that matter) to the effect that the license does not represent a reasonably negotiated value or that the patentee does not have full knowledge of the extent of infringement should be evaluated for possible consideration, appreciating that such statements may be motivated by an intention by one or both of the litigants to have the license excluded as a comparable license in later litigation, irrespective of its actual comparability.

litigation itself, including the parties' desire to avoid legal fees, the financial condition of the parties, the uncertainty and risk of the litigation, and the legal positions of each party.

It should be appreciated, however, that where the presentation of the facts and circumstances of a settlement agreement at trial would significantly increase the time required for the presentation of damages evidence at trial and/or would be confusing to the jury, it may be appropriate for the court to exclude evidence of the settlement agreement at trial. Indeed, in some cases, the discovery of these facts and circumstances may impose a significant burden on one or both parties, such that it may be appropriate for the court to limit or preclude discovery of a settlement agreement during the discovery phase of the case. The admissibility and discoverability of a patent license agreement must be determined on a case-by-case basis in view of various factors. The Working Group does not herein express a view on admissibility or discoverability issues, but rather identifies the factors to be assessed in comparing a settlement license agreement to the hypothetical license at issue that may be useful to the threshold determinations of admissibility and/or discoverability, as well as to the use of settlement agreements in the reasonable royalty analysis in cases where they are admissible. (Relevant to Principle II-8(d)).

Principle II-9: Whether a reasonable royalty should be structured as a running royalty or a lump sum should be explicitly considered in the reasonable royalty analysis.

Comment

It is not a forgone conclusion that every reasonable royalty license must provide for a running royalty. Circumstances may dictate that a reasonable royalty should be paid as a lump sum. A reasonable royalty analysis, therefore, is not complete unless it explicitly considers the proper structure of the royalty payment.

Though royalties are often calculated on a running basis, there may well be factors that suggest the hypothetical negotiation would have led to a lump sum license or to a combination of a lump sum and a running royalty. Regardless of the conclusion, a party advocating a lump sum or a running royalty or some combination of the two must articulate an acceptable basis for its position, just as it must for any other element of a reasonable royalty calculation.

Significant factors that may support a lump sum or a running royalty include: (1) the licensing history of the industry and/or one or both of the parties; (2) the extent to which the financial or competitive situation of the parties favors a lump sum payment as opposed to a continuing running royalty, including: (a) the immediate need of either or both of the parties for capital, (b) each party's perception as to the degree of uncertainty associated with the likely future size of the royalty base, (c) either or both of the parties' risk tolerance as to the possibility that the future revenue stream generated by a running royalty may be unexpectedly large or small, or (d) the competitive burden that would be placed on a licensee by the ongoing payment of a running royalty or on the licensor by the absence of such a payment; (3) whether the benefits of the licensed technology are of a continuing nature, concern a one-time event, or whether the benefits of the technology represent some combination of the two.

With regard to factor (2), issues regarding uncertainty in the market and the risk tolerance of the parties may argue against a lump sum royalty, given that the lump sum structure is often adopted to avoid uncertainty about future sales, and hence, the future royalty stream. Consistent with the Working Group's adoption of the Retrospective Model, the question of uncertainty as it impacts the potential royalty structure should be evaluated in light of all of the information available. Thus, for example, a patented technology that turns out to have enjoyed far greater or far fewer sales during the damages period than what was anticipated at the start of that period might be a good candidate for a lump sum analysis where one or both of the patent owner and infringer were risk averse at the start of the damages period.

This list is not meant to be exclusive, and other factors may be considered in appropriate cases. Likewise, there may be cases where some of the listed factors are not properly considered. As with any other aspect of the reasonable royalty analysis, whether a lump sum royalty is appropriate will turn on the positions of the parties and the value of the invention at the time of the hypothetical negotiation.

If the reasonable royalty is determined to be a lump sum royalty, the size of the lump sum payment should be calculated bearing in mind the principles discussed elsewhere in this Commentary. For example, the difference between a running and a lump sum royalty may be nothing more (or less) than the difference between the present value of the anticipated revenue stream associated with the infringement at the time of the hypothetical negotiation and the present value of the actual revenue stream associated with the infringement during the period of infringement. Where the anticipated revenue stream was highly uncertain and the parties would have favored a lump sum, the lump sum might require adjustment up or down from the appropriately apportioned value of the actual revenue in order to adjust for the uncertainty avoided by the lump sum payment.

III. PRETRIAL PRINCIPLES AND BEST PRACTICES

Principle III: **If and when the court believes that significant questions may exist as to the admissibility of certain damages theories or determinations, then in the appropriate case, the court should consider conducting a hearing after the parties exchange damages contentions to determine: (1) if the parties' damages theories are legally cognizable; (2) if the damages evidence is reliable, relevant and/or admissible; and (3) other disputes relating to damages.**

Comment

Parties in patent litigation move to exclude damages experts' testimony, theories, and evidence in almost every patent case for a variety of reasons, including reliability, applicability, relevance, and prejudice. For example, parties often move to exclude damages experts and/or their theories pursuant to Federal Rule of Evidence 702.¹⁵⁸ Parties also seek to exclude specific damages theories and evidence pursuant to Federal Rules of Evidence 402 and 403.¹⁵⁹

¹⁵⁸ See FED. R. EVID. 702; *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993); *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).

¹⁵⁹ See FED. R. EVID. 402, 403.

Currently, there is no standard procedure or time for courts to consider the reliability and/or admissibility of damages experts' testimony, theories, and evidence. Often, damages expert reports are not disclosed until the end of discovery, after or near summary judgment deadlines. Disputes regarding damages issues therefore are infrequently raised in summary judgment motions. Instead, motions attacking damages theories and evidence are raised in motions *in limine* or in *Daubert* motions. Courts, however, may not want to consider disputes regarding damages issues at the motions *in limine* or *Daubert* stage, which is usually immediately before trial.¹⁶⁰

Furthermore, even if a court does consider motions *in limine*, raising damages issues, particularly on the eve of trial, via these motions may lead to the total exclusion of damages experts, damages theories, and/or evidence on the eve of or during trial. Exclusion of such evidence so late in the process could very likely significantly prejudice a party's ability to present its case at trial.

The parties should propose, and the courts should consider, setting dates for damage contentions to be exchanged by the parties. These contentions should be exchanged in advance of both the close of fact discovery and of the filing of damages expert reports.

In the appropriate case when the courts believe that significant questions may exist as to the admissibility of certain damages theories or determinations, the courts should consider adopting procedures to facilitate the consideration of motions related to the admissibility of damages contentions, theories, and evidence sufficiently in advance of trial such that the parties can account for any adverse rulings before trial. In conjunction with the recommendation that the parties exchange damages contentions, courts should consider conducting a hearing to determine if the parties' damages theories, as detailed in their damages contentions, are legally cognizable, the evidence is reliable, and to resolve any other disputes relating to the damages contentions and theories. Specifically, courts should consider conducting such a hearing, if possible, after any claim construction decision but before the exchange of damages expert reports. In the appropriate case, guidance from the court on whether the parties' damages contentions are legally cognizable prior to the exchange of expert reports will significantly aid the parties and the court.¹⁶¹

As more experience is obtained from the bench and bar seeking to resolve such damages disputes earlier in the litigation, further work should be done by a Sedona Conference Working Group to define and develop how and when the best practices set forth below should be implemented.

BEST PRACTICES

1. *In appropriate cases, when the parties cannot resolve disagreements on the admissibility of certain damages theories, methods, contentions, and evidence, the courts should consider providing guidance to civil litigants*

160 See e.g., *Emcore Corp. v. Optium Corp.*, 2009 U.S. Dist. LEXIS 96305, at *2-3 (W.D. Pa. Oct. 16, 2009) ("A motion *in limine* is not the proper procedural vehicle to raise these issues . . . Optium's Motion does not involve evidentiary rulings, or any other type of issue usually considered on an *in limine* basis. Rather, this Motion resembles a motion for summary judgment . . . The time for filing Motions for Summary Judgment has long passed."); *Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc.*, 2006 U.S. Dist. LEXIS 67562, at *3 (D. Del. Sept. 20, 2006) ("[T]he Court concludes that Defendant's second Motion *In Limine* is akin to a summary judgment motion. In this case, the jury will decide how many infringing sales and offers for sale took place in the United States based on the evidence submitted by the parties. Accordingly, the Court will deny Defendants' motion.")

161 Such procedures may have prevented the damages *Daubert* motions granted in *Apple, Inc. v. Motorola, Inc.*, No. 1:11-cv-8540, 2012 WL 1959560 (N.D. Ill. May 22, 2012), *rev'd*, 2014 WL 1646435 (Fed. Cir. April 25, 2014), where damages expert testimony from both sides was excluded and the case was ultimately dismissed.

in patent cases regarding the procedures for filing and resolving motions related to such disputes.

2. *In such cases, courts can provide schedules that allow for hearings to determine if the parties' damages contentions are legally cognizable, the evidence is reliable, relevant and/or admissible, and other disputes relating to damages theories and contentions.*
3. *In such cases, the attorneys should propose hearings for damages issues after any claim construction decision and before the exchange of expert reports.*
4. *Both parties to a lawsuit should work together prior to the initial case management conference to facilitate the early disclosure of preliminary compensatory damages contentions (PCDCs) and supporting materials.*

Explanation: Federal Rule of Civil Procedure 26(f) requires the parties to meet and confer to discuss the nature and basis of their claims, the possibility for prompt settlement, and the timing of their Rule 26(a)(1) initial disclosures. This meeting must take place no later than twenty-one days before the initial case management conference. One category of information parties must include in their initial disclosure is a “computation of each category of damages and supporting materials.”¹⁶² To allow the parties and the court to gain an early, initial understanding of the compensatory damages theories at issue, any preliminary supporting damages evidence, the potential settlement value of the case, and the scope of potential damages discovery, the parties should work cooperatively to facilitate the exchange of early damages information reasonably in their possession at the time of the Rule 26(f) conference. Further, early damages information may be important to ascertaining whether the scope and expense of discovery is warranted.¹⁶³

Any initial damages calculations and information provided reasonably and in good faith during the PCDCs process are considered preliminary or approximate and may be amended or supplemented. The PCDCs disclosures are not intended to confine a party to the contentions it makes at the outset of the case. It is not unusual for a party in a patent case to learn of additional facts and potential theories of recovery as the case progresses. At the same time, courts should not accept skeletal preliminary compensatory damages disclosures uncritically. Failure to provide good faith damages disclosures and at least “high level” damages discovery at an early stage of the litigation may hinder settlement discussions, and result in unnecessary expenditure of time, money, and judicial resources. Ultimately, in considering early disclosures, the court will need to balance competing considerations on a case-by-case basis.¹⁶⁴

5. *Both parties to a lawsuit should be required to disclose PCDCs and supporting materials concurrently with submission of infringement and invalidity contentions or within a set time after the initial case management conference.*

¹⁶² FED. R. CIV. P. 26(a)(1)(iii).

¹⁶³ See *Eon Corp. IP Holding LLC v. Sensus USA Inc.*, No. C-12-01011 EMC (EDL), 2013 WL 3982994 (N.D. Cal. Mar. 8, 2013).

¹⁶⁴ Compare *Eon Corp. IP Holding LLC*, 2013 WL 3982994 with *Brandywine Commc'ns Techs., LLC v. Cisco Systems, Inc.*, No. C 12-01669 WHA, 2012 WL 5504036 (N.D. Cal. Nov. 13, 2012).

Explanation: Because of the complex nature of patent litigation, parties typically require a fair amount of fact discovery before they understand the other's information sufficiently to formulate even "ball park" damages contentions. Nevertheless, both sides should be required to provide preliminary damages disclosures that are as complete as is reasonably possible, as well as high-level documents in their possession that are likely relevant to a fair assessment of the damages issue. Disclosing initial damages contentions at a relatively early point in the case may allow the parties and the court to assess the value of the case, discuss the scope of potential discovery, make a preliminary evaluation of the possibility of early settlement, and potentially, identify damages issues that should be the subject of early partial summary judgment motions or an evidentiary hearing to test legal theories. Ideally, the parties will conduct their Rule 26(f) conference before the initial case management conference. In those instances, the parties should cooperate to set a time for exchange of PCDCs that is keyed to the submission of infringement contentions and/or the date of the case management conference. By way of example, in jurisdictions that require infringement contentions, the patentee would be expected to submit its PCDCs concurrently with submission of its infringement contentions. The parties accused of infringement would then be expected to submit their PCDCs thirty days thereafter. Alternatively, in jurisdictions that do not mandate infringement contentions, the parties should cooperate to establish a reasonable schedule for exchange of PCDCs subsequent to the case management conference but also keyed to the exchange of infringement contentions through interrogatories or otherwise. In any event, the parties should exchange preliminary damages contention materials and, if appropriate, make them available to the court, subject to appropriate measures to protect confidentiality, including interim protective orders.

6. *The party alleging patent infringement should identify all accused instrumentalities known to it at the time of filing as part of its PCDCs, to the extent not disclosed in any prior infringement contentions.*

Explanation: The notice pleading requirement of the Federal Rules of Civil Procedure under Rule 8(a)(2) requires only "a short and plain statement of the claim showing that the pleader is entitled to relief," in order to "give the defendant fair notice of what the . . . claim is and the grounds upon which it rests."¹⁶⁵ Complaints and counterclaims in most patent cases are worded in a bare-bones fashion, necessitating discovery to flesh out the basis for each party's contentions. To remove any potential ambiguity regarding the preliminary scope of the infringement claims set forth in the pleadings, as part of its PCDCs, the party alleging patent infringement should clearly identify all accused instrumentalities reasonably known to it when it filed its complaint.

7. *As part of its PCDCs, the party alleging patent infringement should identify all theories upon which it bases its potential recovery of compensatory damages and provide a brief explanation of facts supporting those theories.*

165 FED. R. CIV. P. 8(a)(2).

Explanation: Compensatory patent damages traditionally fall into three categories: lost profits, established royalty, and reasonable royalty.

The party asserting patent infringement should identify the theory or combination of theories on which its PCDCs are based. In addition, the party should provide a brief, preliminary explanation of the factual bases that they reasonably know to support the stated theory or theories of recovery.

By way of example only, if a lost profits theory is being asserted, the party asserting patent infringement should explain the reasons and evidence reasonably supporting its contention that it would have made profits “but for” the infringement. One non-exclusive test for proving lost profits is set out in *Panduit Corp. v. Stablin Bros.*¹⁶⁶ If the party asserting infringement relies on the *Panduit* test, then it must provide an explanation of: (1) demand for the patented product during the relevant period; (2) the absence of acceptable non-infringing alternatives to the patented product; (3) manufacturing and marketing capacity to sell the products it claims it could have sold; and (4) if possible, a preliminary calculation of the incremental profit margin. This is just one example of a test for proving entitlement to lost profits, and it is being provided solely to demonstrate the nature and scope of information called for by Best Practice 7.

By way of further example, if a reasonable royalty theory of recovery is being asserted, the party asserting patent infringement should provide a preliminary explanation of the facts supporting its theory of recovery including an identification of the preliminary applicable royalty rate and base, if available, and a brief discussion of all other factors, including a discussion of any relevant *Georgia-Pacific* factors, that may bear on a calculation of a reasonable royalty.¹⁶⁷

The disclosures pursuant to this Best Practice are not intended to confine a party to the damages contentions it makes at the outset of the case. However, failure to provide good faith damages disclosures and at least high-level preliminary compensatory damages contentions at the outset of the litigation may hinder settlement discussions, and result in unnecessary expenditure of time, money and judicial resources.

8. *The party alleging patent infringement should produce to each opposing party, or identify for inspection and copying, all materials supporting its PCDCs theories of recovery.*

Explanation: A party asserting patent infringement should produce copies of all materials reasonably known to it that allegedly support its preliminary compensatory damages theories. This disclosure should be as complete as is reasonably possible and should include high-level documents in the party’s possession concerning its sales and profitability, and those of the industry; market share; comparable license agreements and royalty rates related to the

¹⁶⁶ 575 F.2d 1152, 1166 (6th Cir. 1978).

¹⁶⁷ *Georgia-Pacific*, 318 F. Supp. at 1120.

patent at issue; evidence of demand for the patented features; and basic marketing, pricing, manufacturing, and sales information relating to any products or processes that embody the patented invention, that are licensed under the patent, that compete with, or that are sold with or as a result of products or processes that embody the patented invention or are licensed under the patent.

9. *The party responding to infringement allegations should provide a preliminary response to the PCDCs theories of recovery and, if necessary, identify any applicable alternative theories for calculating compensatory damages in the case.*
 - a. *The accused infringer should identify any reasons why it is locked in to continuing to practice the asserted patent(s), and the time at which it contends that the lock-in occurred.*
 - b. *Any proposed non-infringing alternative designs, or design arounds, must be identified sufficiently early in the fact discovery for the patent holder to take meaningful discovery of its technical and commercial viability.*

Explanation: Although the party responding to a patent infringement allegation may need discovery to fully understand the preliminary damages contentions, it should provide a preliminary high-level response to the theories asserted in the preliminary compensatory damages contentions. By way of example, if the PCDCs theory of recovery is lost profits and an accused infringer has evidence of acceptable non-infringing alternative designs, or design arounds, it should explain why lost profits are unavailable. By way of further example, if the PCDCs theory of recovery is a reasonable royalty and the accused infringer disagrees with the royalty rate because royalty stacking principles apply given marketplace realities, it should discuss this fact in its response.

10. *Each party accused of patent infringement should produce to the party asserting infringement, or identify for inspection and copying, all materials supporting its preliminary response to the PCDCs theories of recovery.*
 - a. *That information should include information about any patents the alleged infringing products actually practice (whether its own or someone else's) and what royalties are paid for licenses to those patents, which is relevant to the royalty stacking inquiry.*
 - b. *The accused infringer should identify, with specificity and clarity, the details of any proposed non-infringing design around alternatives, and identify the persons most knowledgeable about that proffered alternative so that appropriate discovery may be conducted by the patent owner.*

Explanation: A party responding to preliminary damages contentions should produce copies of all materials reasonably known to it that allegedly support its response to the preliminary compensatory damages theories. This disclosure should be as complete as is reasonably possible and should include high-level documents in the party's possession concerning license agreements and royalty rates that relate to the accused product or process; basic

marketing, pricing and sales information relating to the accused products; any non-infringing design around alternatives; and any information that otherwise may be relied upon to define the royalty rate or base.

11. *The parties should establish a date for exchange of final damages contentions that are to be subject to further amendment or supplementation only for good cause.*

Explanation: The preparation of final damages contentions may require significant factual development and may entail involvement of technical and damages experts at considerable expense. Nevertheless, such contentions may be important in focusing and narrowing the damages issues. Ultimately, the timing of exchange of final damages contentions should be set in relation to the cutoff date(s) for fact and expert discovery and for the exchange of final infringement and invalidity contentions. Preferably, the exchange of final damages contentions should occur before the inception of expert discovery but after fact discovery related to damages is sufficiently well developed, and ordinarily after exchange of final infringement and invalidity contentions. The parties would still have an opportunity to amend or supplement the contentions for good cause.

12. *Exchange of PCDCs materials should not be withheld on the basis of confidentiality. The parties should meet and confer in good faith to agree on the form of a suitable protective order.*

Explanation: Disclosure of PCDCs materials should not ordinarily be withheld on the basis of confidentiality. The parties should meet and confer in good faith to agree on the form of a protective order well in advance of the time for the first preliminary damages contention disclosure.

13. *If the parties cannot agree on the form of a protective order sufficiently providing for the protection of financial, licensing, and other confidential damages-related information, and the court where the case is pending does not have a standard protective order for patent cases, a party may apply to the court for a protective order under Fed. R. Civ. P. 26(c).*

Explanation: Where applicable, the protective order authorized by the local rules of the court should govern the disclosure of early compensatory damages contention materials unless the court enters a different protective order *sua sponte* or on motion by a party.

14. *If there is no protective order in place when the PCDCs materials are due, the parties should exchange those materials for "OUTSIDE ATTORNEYS' EYES ONLY." If one of the parties refuses to exchange on that basis, the other can make a motion as outlined above.*

Explanation: Because early disclosure of preliminary compensatory damages theories and support may facilitate settlement and help shape the scope of discovery in the case, the pendency of protective order issues before the court should not delay disclosure of PCDCs materials. If there is no standard protective order set forth by local rule of court, and if the court has not entered a protective order in the case, the parties should agree to exchange the PCDCs materials on an "outside attorneys' eyes only" basis.

15. *The parties should encourage their damages experts to take care to exclude from their reasonable royalty determination any hold-up effects that may result from a valuation performed after the relevant lock-in date. The reasonable royalty analysis should assign the reasonable royalty value prior to the relevant lock-in date. Upon the filing of a Daubert motion challenging the reasonable royalty methodology, the court should explicitly consider whether lock-in/hold-up effects were properly accounted for in the challenged methodology.*
16. *The parties should encourage their damages experts to affirmatively address the issue of royalty stacking in their reports and explain what information they have considered to address this issue.*
17. *Any proposed non-infringing design around alternatives must be subject to the same level of technical expert analysis and judicial scrutiny in the matter that is afforded to the infringement and/or non-infringement analyses.*
 - a. *If the patent holder seeks to challenge introduction of evidence and/or argument regarding the proposed non-infringing design around, it should file a motion with the district court sufficiently in advance of trial. The court should allow the introduction of the proposed alternative only if the accused infringer has shown, by a preponderance of the evidence, that the alternative design does not infringe the asserted claims of the patent(s) in suit.*
 - b. *To the extent the proffered design around is viewed as infringing or unacceptable or both by the patent owner, if the patent owner wishes to prove this at trial, then during expert discovery, the patent owner should produce expert reports from appropriately qualified experts to explain, with specificity and clarity, the reasons why the proffered design around is either infringing or unacceptable.*
 - c. *In response, the accused infringer's expert should provide expert reports from appropriately qualified experts to explain, with specificity and clarity, the reasons why the proffered design around is non-infringing and acceptable.*

IV. TRIAL PRINCIPLES AND BEST PRACTICES

Principle IV-1: Courts can assist in streamlining the presentation of damages evidence at trial to ensure that: (1) damages theories are tied to the specific facts of the case, and that damages experts use reliable methodologies; (2) the entire market value rule is applied only when appropriate; and (3) the comparability of license agreements is rigorously addressed.

Comment

In appropriate cases, courts can rule before trial on whether parties' damages theories are legally cognizable and, in appropriate circumstances, allow parties the opportunity for adjustment should they need to modify their theories.¹⁶⁸ Absent a showing

¹⁶⁸ See, e.g., *Oracle Am., Inc. v. Google Inc.*, 798 F. Supp. 2d 1111 (N.D. Cal. 2011); *Lucent Techs., Inc. v. Microsoft Corp.*, No. 07- CV-2000 H(CAB), 2011 WL 2728317 (S.D. Cal. July 13, 2011).

of good cause (including lack of prejudice to the opposing party), the presumption should be that a party may not modify its damages analysis after some or all of it has been excluded. Courts should be mindful that the patent holder must maintain its burden of proof to establish a reasonable royalty. Some courts have taken the position that, absent an intervening change in the law that affects the admissibility of an expert's damages theory or analysis, it is unfair to give a party that has overreached an opportunity to modify its theories. Yet, as the following examples illustrate, courts can exercise their discretion and allow supplementation.

1. Three months before trial, a judge in the Northern District of California granted the defendant's motion to strike most of the plaintiff's damages expert report in which the expert opined that a royalty "could be as much as \$6.1 billion."¹⁶⁹ The court noted numerous ways in which the opinion was not based on sufficient facts, which included the expert equating the invention with the entirety of Java and Android and his reliance on a mathematical model – unrelated to the specific facts of the case – under which the patentee would be awarded half of the defendant's profits.¹⁷⁰

The court had specifically requested early damages reports, so it would have time to vet the parties' analyses and allow them to adjust their final reports accordingly. After chastising the plaintiff for overreaching in multiple ways, the court cautioned that the next report would be "for keeps" and that the plaintiff should take care to rectify the deficiencies, or the expert would be excluded altogether.¹⁷¹ The court suggested that the plaintiff start its damages analysis with the last offer that was made during negotiations, \$100 million, several years before suit.¹⁷²

2. A judge in the Southern District of California precluded a plaintiff from presenting some of its damages evidence to the jury. The plaintiff was precluded from presenting its theory that the parties would have agreed upon a royalty of \$70 million, which was halfway between the defendant's projected royalty of \$0 and the plaintiff's projected royalty of \$138.7 million. The plaintiff's expert had not, in his single-page analysis, explained why the parties would meet halfway, instead of agreeing upon any other number in that same range.¹⁷³ Similarly, the plaintiff was precluded from presenting evidence on the entire market value, where the plaintiff's expert had not properly apportioned between the accused product's patented and unpatented features.¹⁷⁴ The court reserved the right, however, to revisit the ruling at trial, should the plaintiff meaningfully apportion the per unit price of the accused product.¹⁷⁵ Finally, the plaintiff was allowed to proceed to the jury under a "business realities" approach, wherein it was hypothesized – based on a host of factors – that it would have been unwilling to accept less than \$65–75 million as a lump sum royalty. The court noted, however, that the plaintiff would have to prove that this measure did not violate the entire market value rule, and that the factual bases for the expert's calculations were credible.¹⁷⁶

169 *See Oracle*, 798 F. Supp. 2d at 1114.

170 *Id.* at 1115–16, 1119–21.

171 *Id.* at 1121–22.

172 *Id.* at 1121.

173 *See Lucent*, 2011 WL 2728317, at *9.

174 *Id.* at *5–7.

175 *Id.* at 7.

176 *Id.* at *7–8.

Principle IV-2: A significant amount of trial time should be dedicated to the damages portion of a patent case.

Comment

Courts can ensure that the time allocated for trial is sufficient to permit both sides to fairly address all of the issues to be tried, including all damages issues, in light of the nature and complexity of those issues and the scope of testimony or other evidence needed to address them. Courts faced with busy dockets are, with increasing frequency, ordering timed trials. In light of changes in patent damages law at the Federal Circuit, it should be noted that while two hours for a damages case may have been appropriate in years past, it may not be sufficient today. Depending on the jurisdiction, patent litigants may find themselves in the position of having to put all of their evidence in during a timed trial. Given the recent, increased scrutiny on proof of damages, however, parties cannot afford to skimp on their presentation of damages evidence. In particular, damages experts must thoroughly explain their methodologies, show the evidence they considered, and demonstrate how the evidence impacted their conclusions.

BEST PRACTICES

1. *At or before trial, the parties, with guidance from the court, should determine a fair amount of time for the damages portions of the case; in cases in which there is a concern about the amount of time that a judge will allow for trial, parties should consider reaching agreement as to the admissibility of evidence summarizing an expert's testimony.*
2. *Parties might consider stipulating to the admissibility of summaries under Rule 1006 of the Federal Rules of Evidence,¹⁷⁷ including even summaries that include an expert's calculations, demonstrating the mathematical basis for the opinion.*

Principle IV-3: Bifurcation of a patent damages trial from a patent liability trial may be appropriate.

Principle IV-3(a): In cases involving a single defendant and a single patent, bifurcation of damages may not be appropriate given the relative lack of complexity in the case, potential overlap in proof on various liability and damages issues, and the risk of prejudice to the patentee if infringement continues unabated throughout the time that it takes to try both phases of a case.

Principle IV-3(b): In cases involving multiple defendants, multiple patents and multiple accused products, or those involving particularly complex damages theories, or those in which the courts order timed trials, bifurcation of damages may be appropriate to avoid juror confusion and unnecessary expense.

Principle IV-3(c): If a court decides to completely bifurcate liability discovery and trial from damages

discovery and trial, it should consider also allowing time for an appeal to the Federal Circuit between trials.

Comment

Single-defendant, single-patent cases are unlikely to be bifurcated because they are often relatively straightforward and simple.¹⁷⁸ Proceedings in district court should be administered to be “just, speedy, and inexpensive.”¹⁷⁹ By contrast, the complexity of the evidence on both liability and damages may be overwhelming to a jury in multi-defendant, multi-patent cases.¹⁸⁰

In simple cases, some courts have considered the patentee’s chances of success when deciding whether bifurcation would be more efficient. In a case in the Eastern District of Wisconsin, for example, the court denied the defendant’s motion to bifurcate because it believed the plaintiff was likely to succeed and that a second trial would then be necessary.¹⁸¹ Similarly, in the District of Delaware, the court denied a motion to bifurcate because the defendant had not demonstrated that its “probability of prevailing in its infringement defense [was] incontrovertibly greater than” the patentee’s.¹⁸²

Other courts, however, have more routinely bifurcated cases, taking the view that in all but exceptional patent cases:

[T]he burden imposed on a jury in a patent trial is extraordinary. More specifically, juries are tasked with resolving complex technical issues regarding infringement and invalidity, many times with respect to multiple patents and/or multiple prior art references. Absent bifurcation, jurors then are expected to understand the commercial complexities of the relevant market (or, even more impenetrable, the commercial complexities of the hypothetical market) in order to determine the economic consequences of their liability decisions.¹⁸³

Courts also consider whether evidence related to liability impacts a determination on damages. For example, when an accused infringer mounts a validity challenge under 35 U.S.C. § 103, the patentee may wish to present evidence of secondary considerations of non-obviousness, such as the commercial success of products that practice the patent, the failure of others, and a long-felt need in the industry for the patented invention. This same evidence bears on the determination of a reasonable royalty. A patented invention’s commercial success may, for example, reflect the utility and advantages of the invention over old modes or devices (*Georgia-Pacific* factor 9). Additionally, the failure of others and existence of a long-felt need bear on the amount an accused infringer would have been willing to pay for the invention (*Georgia-Pacific* factors 12 and 15).

177 FED. R. EVID. 1006.

178 See, e.g., *Baratta*, 05-cv-60187, slip op. at 9; *Nielsen*, 2010 U.S. Dist. LEXIS 26804 at *5; see also *BASF Catalysts*, 2009 U.S. Dist. LEXIS 16263 at *6. A “single-defendant, single-patent” case is meant to be only an example of a type of case that can be resolved in a single trial. Other types of cases may also be straightforward enough so that bifurcation is not advisable.

179 FED. R. CIV. P. 1.

180 See, e.g., *Wyeth v. Abbott Labs.*, No. 08-230 (JAP), 2010 U.S. Dist. LEXIS 116921, at *5 (D.N.J. Nov. 3, 2010); *Intel Corp. v. Commonwealth Sci. & Indus. Research Organisation*, Civ. Nos. 6:06-cv-552, 6:06-cv-549, 6:06-cv-550, 6:07-cv-204, 6:06-cv-324, 2008 U.S. Dist. LEXIS 103613, at *28 (E.D. Tex. Dec. 23, 2008); *William Reber, LLC v. Samsung Elecs. Am., Inc.*, 220 F.R.D. 533, 535 (N.D. Ill. 2004).

181 See *Kimberly-Clark Worldwide, Inc. v. First Quality Baby Prods., LLC*, No. 09-C-0916, 2010 U.S. Dist. LEXIS 98573, at *6 (E.D. Wis. Sept. 7, 2010).

182 *Masimo Corp. v. Philips Elecs. N. Am. Corp.*, 742 F. Supp. 2d 492, 500–01 (D. Del. 2010).

183 *Dutch Branch*, 2009 U.S. Dist. LEXIS 76006, at *2–3.

Courts are also disinclined to bifurcate when doing so would severely prejudice a patentee by creating unnecessary delay, and when a defendant's principal goal appears to be to slow the proceedings.¹⁸⁴ Extensive motion practice regarding the admissibility of evidence following bifurcation – such as motion practice related to whether certain evidence should be presented to the jury during the liability trial or reserved for the damages trial – may cause excessive delays in a bifurcated case.¹⁸⁵ Additionally, a patentee is prejudiced by the fact that the appellate process is prolonged in bifurcated cases as each trial may be appealed separately. These separate appeals can cause significant delays in reaching finality, since the litigation must be entirely concluded such that nothing is left except to execute the judgment.¹⁸⁶ A patentee is further prejudiced by the fact that these delays allow for intervening judgments which may vacate an earlier liability judgment.¹⁸⁷

Parties seeking bifurcation should be aware that it is not guaranteed even in multi-defendant cases. Courts may take the view that limiting instructions will suffice to prevent any juror confusion.¹⁸⁸ Alternatively, courts may prefer to manage the complexities of a multi-defendant case in unique ways that are tailored to the parties. For example, the court in the Eastern District of Texas denied a bifurcation motion in a consolidated, multi-defendant case involving 124 defendants.¹⁸⁹ Recognizing that this was not a “typical” patent case – and that the district's local patent rules made defending the case prohibitively expensive – the court set an early *Markman* and summary judgment hearing, and stayed all unrelated discovery.¹⁹⁰ The court's rationale for declining to bifurcate damages was based in part on the patentee's stated strategy of seeking early settlements based on an analysis of each defendant's sales and the cost of defense.¹⁹¹ Given this strategy, damages discovery was necessary in order for the parties to be able to “fully and fairly” evaluate the case for settlement purposes.¹⁹²

Defendants seeking bifurcation should take care to consider the ramifications of a final liability determination. If, following such an outcome, the parties do not settle a case, the plaintiff's strategy during the damages trial will likely include multiple references to the defendant as an “infringer.” Such tactics have the potential to put the defendant at a distinct disadvantage.

The examples below illustrate the issues raised by recommendations made in Principle IV-3:

1. A motion to bifurcate was denied by a district court judge who held the view that “damages and liability are not easily compartmentalized.”¹⁹³ Sales and

184 See, e.g., *BASF Catalysts*, 2009 U.S. Dist. LEXIS 16263, at *6; *Baratta*, 05-cv-60187, slip op. at 9.

185 See *BASF*, 2009 U.S. Dist. LEXIS 16263, at *6 (quoting *Trading Technologies Intern., Inc. v. eSpeed, Inc.*, 431 F. Supp. 2d 834, 840 (N.D. Ill. 2006)) (“Given the nature of this case thus far, we would not be surprised if the parties engaged in extensive motion practice wrangling over whether certain pieces of discovery were applicable to the liability case or the willfulness/damages case. Thus we do not think that defendants have carried their burden of establishing that bifurcation of discovery and trial would promote judicial efficiency.”) See also *Baratta*, No. 05-cv-60187, slip op. at 9 (“In particular, the Court is concerned in this case, and in light of the lack of progress that has occurred in the past three and a half years, that bifurcation would serve to further prolong this matter by creating additional discovery periods, additional trials, and additional motions for relief.”).

186 See e.g., *Fresenius USA, Inc. v. Baxter Intern., Inc.*, 721 F.3d 1330, 1341 (Fed. Cir. 2013) (*en banc* rehearing denied); see also, *Robert Bosch*, 719 F.3d at 1305 (holding that the Federal Circuit has jurisdiction to entertain appeals from patent infringement liability determinations when a trial on damages has not yet occurred).

187 *Id.* at 1332 (The Federal Circuit remanded the initial case to the district court to reconsider its damages verdict. While the litigation was pending on remand, the United States Patent Office completed its re-examination proceedings and determined that all of the claims were invalid. The Federal Circuit affirmed the PTO's determination, vacated the district court's judgment, and remanded with instructions to dismiss the case.).

188 See *Lutron Elecs.*, 2010 U.S. Dist. LEXIS 49623, at *6.

189 See *Parallel Networks LLC v. AEO, Inc.*, No. 6:10-cv-00111, Dkt. No. 338, slip op. at 8 (E.D. Tex. Mar. 15, 2011).

190 *Id.* at 6.

191 See *id.* at 4.

192 *Id.* at 8.

193 *Kimberly-Clark*, 2010 U.S. Dist. LEXIS 98573, at *5.

financial information would be considered by the jury in determining whether the patentee had proven “commercial success,” and that same information “is inherently intertwined with damages.”¹⁹⁴

2. In denying a bifurcation motion, a court may consider that an accused infringer’s proof of non-infringement and invalidity has ramifications not only for a liability determination, but also for a determination of whether infringement was willful. The patentee must prove by clear and convincing evidence that an infringer acted despite “an objectively high likelihood that its actions constituted infringement of a valid patent.”¹⁹⁵ The defendant’s actions must have been “objectively reckless.”¹⁹⁶ It is not likely that a defendant will be found to have acted willfully if it has raised a reasonable defense to infringement. For this reason, questions may exist as to whether willfulness is more appropriately tried with liability or damages, and ambivalence over when to hear evidence on willfulness may make a judge more inclined to deny a bifurcation motion.¹⁹⁷

Litigants relying on a retrospective approach to the determination of a reasonable royalty may consider the efficiencies in undertaking damages discovery just once, at a point in time that is late enough to allow them to gather all of the relevant discovery. Litigants should also be mindful, though, that regardless of whether their damages model is prospective or retrospective, a significant delay in taking damages discovery creates the risk that discovery closest in time to the date of first infringement will be lost. However, depending on the situation, it may be preferable to conclude all discovery and then have a staged trial with the same jury rather than different juries.

BEST PRACTICES

1. *Where a case is bifurcated, litigants should consider whether discovery should also be bifurcated in light of their damages theories, or whether it is preferable to conclude all discovery at once.*
2. *If discovery is completed on all issues, bifurcated trials would benefit from having the same jury, whereas in cases where discovery is bifurcated, the trials will be to different juries.*

Principle IV-4: In a typical case, a willfulness allegation should not itself dictate a bifurcation of damages from liability. To the extent possible, where a case is bifurcated, and willfulness is tried after liability is determined, it is preferable to have a staged trial with the same jury rather than different juries.

¹⁹⁴ *Id.*

¹⁹⁵ *In re Seagate Tech., LLC*, 497 F.3d 1360, 1371 (Fed. Cir. 2007).

¹⁹⁶ *Id.*

¹⁹⁷ *See, e.g.*, Kathleen Burdette Shields & Jessica Gan Lee, *The Bifurcation Divide*, LAW360, (Nov. 18, 2009, 10:45 AM), <http://www.law360.com/articles/130440/the-bifurcation-divide>.

Comment

Courts have the authority and discretion to try the issues of liability, willfulness, and damages together or separately.¹⁹⁸ Appeals may be entertained on patent infringement liability determinations when willfulness issues are outstanding and remain undecided.¹⁹⁹

Courts that have refused to bifurcate willfulness from liability have declined to do so because “[m]any of the witnesses and evidence needed to address the willfulness issue are the same as that needed to address the liability issue.”²⁰⁰ Because willfulness is determined from the totality of the circumstances, those courts have concluded it is necessary for a jury to “look at all of the evidence as a whole.”²⁰¹

By contrast, other courts have bifurcated liability from willfulness and damages to avoid juror confusion when there are multiple defendants, which requires inquiry into the state of mind of each of the defendants, as well as into the attendant facts and circumstances.²⁰²

In light of the Principle above, a party may consider seeking bifurcation of willfulness from liability to the extent that it plans to rely upon an “advice of counsel” defense against willfulness, but it does not want to prove up its attorney-client communications during the liability trial.²⁰³

However, because it is preferable to have the same jury determine liability and willfulness, discovery on willfulness should be completed before the liability trial, so the trials can be staged one after another with the same jury. Whether an infringer’s proofs on invalidity and non-infringement are consistent with its pre-infringement opinion of counsel may be probative to the infringer’s good faith.²⁰⁴

Principle IV-5: Jury instructions that are tailored to the case will be more suitable than model jury instructions.

Comment

When a jury is charged with making a reasonable royalty determination, litigants should consider what the jury is told about the *Georgia-Pacific* factors. For example, litigants should ask if the jury should even be aware of certain factors if there is no testimony on those particular factors, because the jury may draw inferences from the absence of testimony on those factors. A better approach in such cases may be to reframe the instructions to ask the jury to focus on the invention, its contribution over the prior art, and the *Georgia-Pacific* factors present in the case.

Accordingly, if the damages expert witnesses only rely upon a subset of *Georgia-Pacific* factors, the jury should only be instructed on those factors. If comparability of licenses is at issue, the jury should be given specific guidance on how to determine

198 *Robert Bosch*, 719 F.3d at 1319–20.

199 *Id.* at 1317.

200 *Robotic Vision Sys., Inc. v. View Eng'g., Inc.*, No. 96-cv-2288, 1997 U.S. Dist. LEXIS 19157 at *8 (C.D. Cal. Nov. 6, 1997).

201 *Real v. Bunn-O-Matic Corp.*, 195 F.R.D. 618, 626 (N.D. Ill. 2000).

202 *See Medpointe Healthcare Inc. v. Hi-Tech Pharmaceutical Co.*, No. 03-5550, 2007 U.S. Dist. LEXIS 4652 at *16–17 (D.N.J. Jan. 22, 2007).

203 *See, e.g., Real*, 195 F.R.D. at 625.

204 *Patent Holding Co. v. TG (USA) Corp.*, No. 96-cv-74721, 1997 U.S. Dist. LEXIS 22571 at *4–5 (E.D. Mich. Dec. 2, 1997).

comparability. Moreover, model jury instructions do not adequately address the entire market value rule and how to determine an appropriate base; thus, courts must craft new instructions based on the particular facts of the case, current case law, and the principles articulated in this paper.

Principle IV-6: Jury verdict forms that are tailored to the case will be more suitable than general verdict forms. Thus, in most cases, the verdict form should ask the jury to determine an amount of damages adequate to compensate for the infringement, on a per patent/per claim basis. Also, special verdict forms may be preferable in cases involving ongoing damages.

Comment

Litigants should be aware of the risks and advantages of different verdict form formats. The jury verdict form should be sufficiently detailed to avoid the need for remand and retrial after appeal. For example, where there are multiple patents, damages should be identified for infringement of each patent and on each claim found infringed so that reversal of validity or infringement of one patent would not require remand and retrial of damages on all patents-in-suit. On the other hand, increased specificity can increase the risk of juror confusion and inconsistent verdicts.

A jury might simply be asked to determine a number adequate to compensate for infringement. In cases in which ongoing infringement is a concern, juries should be asked to determine both the damages base and the applicable royalty rate, but should not be asked to perform the ultimate calculation.

In the alternative, jurors could be presented with special verdict forms where they are asked to make factual determinations, allowing the judge to apply the relevant law. Or, special verdict forms might be drafted to include special interrogatories. For example, in cases of ongoing infringement, the parties may desire that a jury determine whether an ongoing running royalty, or a lump sum payment, is appropriate. In other cases, where one party asserts that the reasonable royalty should take the form of a lump sum, but the parties do not agree to submit the question of future damages to the jury, it may be beneficial to instruct the jury as to the dates covered by the reasonable royalty the jury awards.

In cases involving multiple patents and/or multiple accused products, the parties should consider whether a special verdict form is warranted, to ensure clarity on remand. On the other hand, a patentee may take the approach that it is the defendant's burden to appeal any part of a damages determination that it wishes to challenge on remand.

V. POSTTRIAL PRINCIPLES AND BEST PRACTICES

A. Injunctions

Principle V-1: A patent is a property right and the patentee usually is irreparably harmed if the right to exclude is not enforced.

Comment

The Supreme Court in *eBay, Inc. v. MercExchange, LLC*,²⁰⁵ held that to obtain a permanent injunction, a patentee must demonstrate that: (1) it has suffered an irreparable injury; (2) the remedies available at law are inadequate to compensate for that injury; (3) hardships between the plaintiff and defendant favors an injunction; and (4) the public interest would not be disserved by issuance of an injunction.²⁰⁶

The question of whether there is a rebuttable presumption of irreparable harm was left unanswered in *eBay*, but addressed in two concurring opinions.

Roberts Concurrence

Chief Justice Roberts's concurrence (joined by Justices Scalia and Ginsburg) paralleled the prior view of the Federal Circuit:

[The] 'long tradition of equity practice' [granting injunctive relief upon finding infringement] is not surprising, given the difficulty of protecting a right to *exclude* through monetary remedies that allow an infringer to *use* an invention against the patentee's wishes – a difficulty that often implicates the first two factors of the traditional four-factor test.²⁰⁷

Kennedy Concurrence

Justice Kennedy's concurrence (joined by Justices Stevens, Souter, and Breyer) expressed concern over the Federal Circuit's prior view:

In cases now arising trial courts should bear in mind that in many instances the nature of the patent being enforced and the economic function of the patent holder presents considerations quite unlike earlier cases.²⁰⁸

Justice Kennedy's concurrence specifically called out the following issues a court should consider when deciding whether to issue an injunction:

- a. non-practicing entities (NPEs) ("An industry has developed in which firms use patents not as a basis for producing and selling goods but, instead, primarily for obtaining licensing fees. For these firms, an injunction, and the potentially serious sanctions arising from its violation, can be employed as a

²⁰⁵ 547 U.S. 388 (2006).

²⁰⁶ *Id.* at 391, 393–94.

²⁰⁷ *Id.* at 395.

²⁰⁸ *Id.* at 396.

bargaining tool to charge exorbitant fees to companies that seek to buy licenses to practice the patent.”);²⁰⁹

- b. small patented components of a larger accused device (“When the patented invention is but a small component of the product the companies seek to produce and the threat of an injunction is employed simply for undue leverage in negotiations, legal damages may well be sufficient to compensate for the infringement and an injunction may not serve the public interest.”);²¹⁰ and
- c. business method patents (“In addition injunctive relief may have different consequences for the burgeoning number of patents over business methods, which were not of much economic and legal significance in earlier times. The potential vagueness and suspect validity of some of these patents may affect the calculus under the four-factor test.”).²¹¹

Federal Circuit Reaction

Subsequent to *eBay*, the Federal Circuit has interpreted the Supreme Court’s decision to have removed the presumption of irreparable harm. However, consistent with Chief Justice Roberts’s concurrence, the Federal Circuit in *Robert Bosch, LLC v. Pylon Manufacturing Corp.* clarified that although “*eBay* jettisoned the presumption of irreparable harm as it applies to determining the appropriateness of injunctive relief,” the right to exclude, fundamental to patent law, should not be ignored.²¹² Specifically, the Federal Circuit stated:

[a]lthough *eBay* abolishes our general rule that an injunction normally will issue when a patent is found to have been valid and infringed, it does not swing the pendulum in the opposite direction. In other words, even though a successful patent infringement plaintiff can no longer rely on presumptions or other short-cuts to support a request for a permanent injunction, it does not follow that courts should entirely ignore the fundamental nature of patents as property rights granting the owner the right to exclude. Indeed, this right has its roots in the Constitution, as the Intellectual Property Clause of the Constitution itself refers to inventors’ “exclusive Right to their respective . . . Discoveries.” U.S. Const. Art. I, § 8, cl. 8 (emphasis added).²¹³

Similarly, in *Edwards Lifesciences AG v. Corevalve, Inc.*, the Federal Circuit reiterated that “[t]he Court in *eBay* did not hold that there is a presumption *against* exclusivity on successful infringement litigation.”²¹⁴ Rather, “[a]bsent adverse equitable considerations, the winner of a judgment of validity and infringement may normally expect to regain the exclusivity that was lost with the infringement.”²¹⁵

209 *Id.* (citations omitted).

210 *Id.* at 396–97.

211 *Id.* at 397.

212 *Robert Bosch*, 659 F.3d at 1149.

213 *Id.*

214 *Edwards Lifesciences AG v. Corevalve, Inc.*, 699 F.3d 1305 (Fed. Cir. 2012) (emphasis added).

215 *Id.* at 1314.

The Federal Circuit has reiterated that there is neither a presumption for nor against an injunction. Whether an injunction should issue depends on the facts of the case and a proper weighing of the equitable considerations.

Thus, a district court must consider the patentee's right to exclude in determining whether an injunction is an appropriate remedy. However, the district court must weigh the equities as set out by the Supreme Court in *eBay* and may not presume irreparable harm or the inadequacy of monetary relief.

Comparison to ITC

The Federal Circuit has held that the *eBay* decision does not apply to exclusion orders in patent cases before the International Trade Commission. In *Spansion, Inc. v. International Trade Commission*, the Federal Circuit found that the applicable statute requires the Commission to issue an exclusion order upon finding a violation under Section 337, noting that “[t]he legislative history of the amendments to Section 337 indicates that Congress intended injunctive relief to be the normal remedy for a Section 337 violation and that a showing of irreparable harm is not required to receive such injunctive relief.”²¹⁶ Rather, the statute requires consideration of specific public interest factors that include: the public health and welfare; competitive conditions in the United States economy; the production of like or directly competitive articles in the United States; and United States consumers.²¹⁷

Stay Pending Appeal

Where appropriate, a permanent injunction may be stayed pending appeal. A court may issue such a stay pursuant to FRCP 62(c), which states that “[w]hile an appeal is pending from an interlocutory order or final judgment that grants, dissolves, or denies an injunction, the court may suspend, modify, restore, or grant an injunction on terms for bond or other terms that secure the opposing party’s rights.”²¹⁸ A stay of an injunction pending appeal may be obtained at the district court or the Federal Circuit.²¹⁹

In determining whether to stay an injunction pending appeal, the district court and the Federal Circuit apply the same test, by considering the following four factors:

1. whether the stay applicant has made a strong showing that he is likely to succeed on the merits;
2. whether the applicant will be irreparably injured absent a stay;
3. whether issuance of the stay will substantially injure the other parties interested in the proceeding; and
4. where the public interest lies.²²⁰

²¹⁶ *Spansion, Inc. v. International Trade Comm'n*, 629 F.3d 1331, 1358 (citing to 19 U.S.C. § 1337(d)(1)).

²¹⁷ *Id.*

²¹⁸ FED. R. CIV. P. 62(c).

²¹⁹ FED. R. APP. P. 8(a)(1)–(2).

²²⁰ *Standard Havens Prods., Inc. v. Gencor Indus., Inc.*, 897 F.2d 511, 512 (Fed. Cir. 1990) (quoting *Hilton v. Braunskill*, 481 U.S. 770, 776 (1987)).

Thus, for example, in a case in which the claim construction or other issues on the merits were not clearly in favor of one party, the presiding district court that enters an injunction in favor of a patent owner could stay the injunction pending resolution of the appeal. Under those circumstances, any settlement negotiations will be based on the parties' evaluation of the strength of their respective positions on appeal, not on the *in terrorem* effect of the threat of being excluded from the market before the appeal can be decided. As noted below, the court did precisely that in *Smith & Nephew, Inc. v. Arthrex, Inc.*²²¹

As an alternative to staying an injunction pending appeal, another option available in appropriate circumstances is for a court to issue a permanent injunction, but provide for a sunset period for the defendant to implement a non-infringing alternative.²²² In these circumstances, the patentee is typically compensated for the continued use of its patent through the payment of sunset royalties.²²³

The following cases exemplify the use of a stay of injunction pending appeal:

a. *i4i v. Microsoft*

While a stay pending appeal was denied at the district court,²²⁴ a stay pending appeal was granted by the Federal Circuit.²²⁵ Thereafter, the Federal Circuit affirmed the finding of infringement and reinstated the injunction but modified the effective date of the injunction.²²⁶

b. *Verizon v. Vonage*

The Federal Circuit granted a stay pending appeal²²⁷ after the district court stayed the injunction pending appeal with respect to present or existing customers, provided Vonage escrowed the 5.5% royalty quarterly.²²⁸ The Federal Circuit later affirmed the injunction as to two patents but vacated the judgment of infringement with respect to a third patent, and remanded for a new trial.²²⁹

c. *NTP v. RIM*

The district court granted a stay of the permanent injunction pending appeal.²³⁰ The court granted the plaintiff's motion for permanent injunction, finding that:

221 No. 2:07-cv-335, 2010 WL 2522428 (E.D. Tex. June 18, 2010).

222 *See, e.g., Broadcom Corp. v. Emulex Corp.*, 732 F.3d 1325, 1339 (Fed. Cir. 2013) (“[T]he district court’s selection of an eighteen month sunset period was not an abuse of discretion. The eighteen months allowed for time to remove the infringing product from the market without causing significant downstream disturbance for OEMs and consumers. And the eighteen-month period is a compromise between the wide range of time estimates in the record relating to the design process and product qualification.”).

223 *See e.g., Active Video Networks, Inc. v. Verizon Commc’ns, Inc.*, 694 F.3d 1312, 1342–43 (Fed. Cir. 2012) (affirming district court’s imposition of a sunset royalty).

224 *i4i Ltd. P’ship v. Microsoft Corp.*, 670 F. Supp. 2d 568, 602–03 (E.D. Tex. 2009).

225 *i4i Ltd. P’ship v. Microsoft Corp.*, 343 Fed. Appx. 619 (Fed. Cir. 2009).

226 *i4i Ltd.*, 598 F.3d at 863–64.

227 *Verizon Servs. Corp. v. Vonage Holdings Corp.*, 228 Fed. Appx. 986 (Fed. Cir. Apr. 24, 2007).

228 *Verizon Servs. Corp. v. Vonage Holdings Corp.*, No. 1:06-CV-682 (E.D. Va. Apr. 12, 2007) (D.I. 549).

229 *Verizon Servs. Corp. v. Vonage Holdings Corp.*, 503 F.3d 1295 (Fed. Cir. 2007).

230 *NTP, Inc. v. Research in Motion, Ltd.*, No. 3:01CV767, 2003 WL 23100881 (E.D. Va. Aug. 5, 2003).

1. NTP will be face [sic] irreparable harm if an injunction is not issued;
2. NTP has no adequate remedy at law to address future infringing sales;
3. an injunction in this case is in the public interest as it promotes protection of the rights gained through the patent process; and
4. the balance of hardships between NTP as the holder of the patents-in-suit, and Research in Motion, Ltd. (“RIM”) as the infringing party, weighs more heavily towards NTP.²³¹

The above order notwithstanding, however, the court also granted the defendant’s motion for a stay, finding that:

1. RIM will be irreparably injured absent a stay of the permanent injunction;
2. the issuance of the stay will not substantially injure NTP; and
3. issuance of the stay is in the public interest, as the public has a demonstrated and increasing use of the products and services involved in this litigation.²³²

The damage award and injunction were vacated on appeal.²³³

d. Smith & Nephew v. Arthrex

After considering the four-factor test reiterated in *eBay*, the district court granted Smith & Nephew’s motion for a permanent injunction, but stayed the permanent injunction pending appeal, finding that “the facts and legal issues of this case are particularly close on the issue of infringement.”²³⁴ The infringement decision was reversed on appeal.²³⁵

BEST PRACTICES

Guidance for various scenarios with respect to Principle V-1 is provided below, although each case should be decided on its specific facts:

1. *The presence of direct competitors presents the strongest case for the court to enter an injunction and deny a stay. The patent owner has a clear economic market interest in excluding the competing infringer from the marketplace, regardless of whether the patent owner practices the patented invention. This also presents the case where it is*

²³¹ *Id.* at *1.

²³² *Id.* at *2.

²³³ *NTP, Inc. v. Research in Motion, Ltd.*, 418 F.3d 1282, 1325 (Fed. Cir. 2005).

²³⁴ *Smith & Nephew, Inc. v. Arthrex, Inc.*, No. 2:07-cv-335, 2010 WL 2522428, at *4 (E.D. Tex. June 18, 2010).

²³⁵ *Smith & Nephew, Inc. v. Arthrex, Inc.*, 453 Fed. Appx. 977, 981 (Fed. Cir. 2011).

least likely that the patent owner is using the injunction to leverage a premium unrelated to the damages which could be suffered pending appeal.

2. *A patent owner's practice of licensing the patent widely to whomever has requested a license presents a strong case for denial of the injunction.*
3. *In cases where the patentee practices the patent, but the litigants are not competitors, depending on the circumstances of the matter, the patentee may not be entitled to an injunction.*²³⁶
4. *In cases where the patentee does not practice the patent, but licenses it to an exclusive licensee who practices the patent and is a direct competitor with the infringer, the patentee may be entitled to an injunction. Similarly, in some circumstances where there was extremely limited licensing to licensees who practice the patent and are direct competitors with the infringer, the patentee may be entitled to an injunction.*
5. *In cases where the patentee does not practice the patent and licenses it to one or very few licensees who do not practice the patent, the patentee may not be entitled to an injunction.*
6. *An injunction should generally not be entered if the patentee is asserting that the patent is essential to a standard and the patentee has made a fair, reasonable, and non-discriminatory (FRAND) commitment.*
7. *A stay of the injunction pending appeal should be considered as part of the overall injunction analysis. For example, the district court should consider whether a stay of the injunction will cause irreparable harm to a patentee-competitor.*²³⁷

B. Alternatives to Injunctions

Principle V-2: If an injunction is not available, then ongoing royalties may be available.

Comment

The Federal Circuit has held that “[t]he award of an ongoing royalty instead of a permanent injunction to compensate for future infringement is appropriate in some cases.”²³⁸

For example, in *Edwards Lifesciences AG v. Corevalve, Inc.*, the Federal Circuit outlined that:

Precedent illustrates the variety of equitable considerations, and responsive equitable remedy in patent cases; for example, the grant of a

²³⁶ A common scenario where this occurs is when the patent covers only a component of a larger product. In that case, the importance of the component to the overall product will be an important factor to consider. If the invention is a critical component, a non-competitor patentee may still be entitled to an injunction. An additional example is when a patent covers a method of manufacture or chemical/biological intermediate that is capable of producing different end products, a particular product with multiple end uses, or results in significant production cost savings. Depending on the circumstances, an injunction might be appropriate, particularly if the patentee has exclusively licensed a competitor of the infringer, or if the infringement impacts the patentee's ability to grant field-of-use licenses or to comply with a requirements contract.

²³⁷ See *Robert Bosch*, 659 F.3d at 1151 (“[t]he existence of a two-player market may well serve as a substantial ground for granting an injunction – for example, because it creates an inference that an infringing sale amounts to a lost sale for the patentee.”); *i4i v. Microsoft*, 598 F.3d 831, 861 (Fed. Cir. 2010).

²³⁸ *Bard Peripheral Vascular, Inc. v. W.L. Gore & Assocs., Inc.*, 670 F.3d 1171, 1192 (Fed. Cir. 2012) (citing *Paice*, 504 F.3d 1314).

royalty-bearing license instead of imposing an injunction in situations where the patentee would experience no competitive injury, as in *ActiveVideo Networks, Inc. v. Verizon Communications, Inc.*, or where there is an overriding public interest in continued provision of the infringing product, as in *Bard Peripheral Vascular, Inc. v. W.L. Gore & Assocs., Inc.*, where the Gore vascular graft materials were not available from the successful patentee Bard. Another form of equitable response is illustrated in *Broadcom Corp. v. Qualcomm Inc.*, where the court postponed the effective date of an injunction for twenty months, to relieve hardship on the infringer.²³⁹

A judgment of an ongoing royalty for post-verdict infringement will only be granted where equitable relief, in the form of a permanent injunction, is not granted.

Recent Federal Circuit case law has explored the tension between awarding a patentee damages as opposed to an injunction. In *Paice LLC v. Toyota Motor Corp.*, the court stated that in certain cases, awarding the patentee with an ongoing royalty, rather than an injunction, may be the appropriate course of action.²⁴⁰ The Federal Circuit stated that the text of Section 283, that “empowers ‘courts . . . [to] grant injunctions in accordance with the principles of equity . . . on such terms as the court deems reasonable,’ leaves no doubt that Congress did not intend to statutorily entitle patentees to a jury trial for the purposes of awarding relief thereunder.”²⁴¹ The court accepted Paice’s argument that “the determination of damages is a legal question which carries a Seventh Amendment right to a jury trial,” but qualified this statement by stating that “not all monetary relief is properly characterized as ‘damages.’”²⁴² Several years later, the court addressed the same issues in *Telecordia Techs., Inc. v. Cisco Sys., Inc.*,²⁴³ and affirmed the views stated in *Paice*.

According to the Federal Circuit in *Paice*, the Seventh Amendment does not apply to an ongoing royalty determination because the court can determine the mandatory royalty as an *equitable* alternative to an injunction.²⁴⁴ This holding appears to present an inconsistent result in the following scenario: if a patentee sues only for back damages, never asking for an injunction or a forward royalty, and then sues every six months for damages, the patentee would be entitled to a jury trial in each of those cases. This practice would be highly inefficient, both for the patentee and the courts. While it is unclear why the result should be different when the patentee acts more efficiently by suing for both back damages and a forward royalty at once, *Paice* is the current law, although the Federal Circuit did not explain why patent damages should be treated differently than any other continuing tort.

There are valid arguments on both sides of the Federal Circuit’s decision in *Paice*. For that reason the Working Group states that ongoing royalties “may be available” as an alternative to an injunction. Opponents of the Federal Circuit’s decision in *Paice* provided the following arguments: The Patent Act provides for the award of damages to the patentee upon a finding of infringement in Section 284.²⁴⁵ Title 35 also provides that in appropriate

²³⁹ *Edwards Lifesciences*, 699 F.3d at 1315.

²⁴⁰ *Paice*, 504 F.3d at 1314. The Federal Circuit also stated that should the district court decide that an ongoing royalty is the more appropriate remedy, the district court has the discretion to permit the parties to negotiate a license agreement themselves before imposing an ongoing royalty on the parties. *Id.* at 1315. In this way, the district court attempts to most closely approximate the hypothetical negotiation described *supra* in Chapter II.

²⁴¹ *Id.* at 1293, n.16 (emphases added).

²⁴² *Id.* at 1316.

²⁴³ *Telecordia Techs., Inc. v. Cisco Sys., Inc.*, 612 F.3d 1365, 1378–79 (Fed. Cir. 2010).

²⁴⁴ *Id.*

²⁴⁵ See *supra* Chapter II.

circumstances, a court *may* grant an injunction to a prevailing party “in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable.”²⁴⁶ Section 283 authorizes a court to grant an injunction if the circumstances warrant it, but does *not* authorize monetary damages as an equitable alternative to be determined *by the court* if it declines to enter an injunction. Further, Section 284 does not authorize monetary damages without a jury trial. The Seventh Amendment supports this view as well. It is well-established that if an issue was tried before a jury at common law at the time this country was founded, or is analogous to an issue that was so tried, the Seventh Amendment mandates a jury trial on that issue unless the parties waive this right.²⁴⁷ Under *Markman* and related precedent, the Federal Circuit should conclude that the assessment of monetary damages in patent cases is analogous to issues tried before a jury at common law, thus requiring a jury determination in cases today. Accordingly, contrary to the holding in *Paice*, both Section 284 and the Seventh Amendment require a jury to resolve the royalty rate applicable to post-verdict infringement. It seems most logical that if an action for past infringement and past damages entitles a patentee to a jury trial, then a judgment that the continued infringement (i.e., an ongoing tort) results in additional damages and the amount of those damages should be treated the same way. The reasoning in *Paice* does not sufficiently justify why an ongoing royalty ceases to be “damages” or is an equitable issue.

Principle V-3: An ongoing royalty should fairly compensate the patent holder for the ongoing use made by the infringer of the patented invention and should be determined by considering what fully informed and reasonable persons in the position of the patent owner (or owners throughout the period of infringement) and the infringer would agree to at the time of trial as a fair price for the license, from the time of trial through the expiration of the patent, taking into account all relevant facts and circumstances occurring before or during that period.

Comment

As discussed above with respect to a reasonable royalty for past damages, an ongoing royalty for future damages should fairly compensate the patent holder for the actual use made by the infringer of the patented invention. This view is consistent with the statutory mandate that damages should be “adequate to compensate for infringement.” An ongoing royalty awarded at the conclusion of a trial is in lieu of the patent holder filing a later suit(s) for damages for the ongoing use. Of course, if the jury awarded a fully paid up lump sum amount, then no ongoing royalty would be owed.

BEST PRACTICES

1. *Rather than simply applying the pre-verdict royalty rate to post-verdict conduct, specific evidence should be presented as to a post-verdict royalty.*
2. *Courts should adhere to the following principles in addressing a post-verdict royalty:*

²⁴⁶ 35 U.S.C. § 283.

²⁴⁷ *See Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 377 (1996) (holding that claim construction is a matter of law for the court to determine and distinguishing claim construction from issues historically left for the jury).

- a. *Forward damages should start at the conclusion of the trial.*
 - b. *The royalty should fairly compensate the patent holder for the ongoing use made by the infringer of the patented invention.*
3. *Courts should also consider the following with respect to determining a post-verdict royalty:*
- a. *How does the change in bargaining positions and/or economic conditions affect the royalty rate?²⁴⁸*
 - b. *Is the infringer now deemed to be willful?*
 - c. *Is the “willing licensee” and “willing licensor” paradigm still appropriate?²⁴⁹*
 - d. *Should the Georgia-Pacific factors be applied and, if so, which ones? Should the focus be on factors that may have changed from the original hypothetical negotiation, such as the existence of design around products, the value of the technology, and the willfulness of the post-verdict infringement?*
 - e. *Should there be a single ongoing royalty rate, or a varying rate (e.g., one that increases over time)?*
 - f. *Should different industries and/or technologies be treated differently? This consideration stems from the idea that what makes sense in one technical field might make little sense in another technical field, and a “one size fits all” approach is not good practice.*
4. *The following timing may be used to hear evidence on the post-verdict royalty:*
- a. *After the motion for permanent injunction is denied;*
 - b. *During trial, while the jury is empaneled (e.g., presenting the issue of an ongoing royalty rate to the jury, notwithstanding that the patentee intends to seek an injunction); or*
 - c. *Through the filing of a separate complaint.*

C. Attorneys’ Fees and Fee Shifting²⁵⁰

Principle V-4: Pursuant to 35 U.S.C. § 285, where a party to a patent lawsuit improperly initiates or maintains one or more claims or defenses, an award of attorneys’ fees is presumptively appropriate. Attorneys’ fees may be assessed against the party and/or its counsel as circumstances warrant.

²⁴⁸ See *Amado* at 1362.

²⁴⁹ See *Soverain Software*, 899 F. Supp. 2d at 589–90 (applying a 2.5x enhancement to the jury’s implied royalty rate for ongoing royalties).

²⁵⁰ This paper does not address willful patent infringement or the potential enhanced damages and attorneys’ fees recoverable after such a finding.

Comment

In April 2014, in addressing the standard for deciding whether a case is “exceptional” for the purpose of awarding attorneys’ fees under 35 U.S.C. § 285, the Supreme Court lowered the bar from the former “objectively and subjectively baseless” standard to one that covers litigation practices that “stand[] out from others with respect to the substantive strength of a party’s litigating position (considering both the governing law and facts of the case) or the unreasonable manner in which the case was litigated.”²⁵¹ Reference in this section to claims or defenses that are “improperly” asserted or maintained should be understood to refer to the currently applicable Supreme Court standard.

Whether or not to award fees is typically within the discretion of the district court. The Working Group’s formulation that an award of fees is “presumptively appropriate” is meant to state the Working Group’s view that it is an abuse of the court’s discretion to refrain from awarding fees where litigation is improperly pursued.

We have noted that such fee shifting is appropriate where “one or more” claims or defenses are improper. This is meant to recognize that courts have the flexibility to sanction conduct, even where some claims or defenses in a case may be appropriately pursued. We expect that litigants and the courts will be self-regulating in seeking and applying fee shifting only to improper claims and defenses that *have a material impact* on the scope and cost of a litigation, and not to pursue sanctions with respect to each failed claim or defense without regard to whether it has had any material impact on the scope or burden of the litigation.

Principle V-5: A claim or defense is improper at whatever point in time it becomes the case that:

- 1. The patentee and/or its counsel had actual knowledge that (or were willfully blind as to the fact that) the asserted claim is either not valid, not infringed, or not enforceable;**
- 2. The accused infringer and/or its counsel had actual knowledge that (or were willfully blind as to the fact that), contrary to an asserted defense, the asserted claim is valid, infringed, or enforceable;**
- 3. In the case of a party to a patent lawsuit, a reasonable person in the position of the patentee and advised by competent counsel would understand that the pursuit of a claim or defense is without merit; or**
- 4. In the case of the party’s counsel, competent counsel in the position of the party’s counsel would understand that the pursuit of a claim or defense is without merit.**

²⁵¹ See *Octane Fitness*, 134 S.Ct. at 1756; see also *Highmark Inc. v. Allcare Health Mgmt. Sys.*, 134 S.Ct. 1744, 1746 (2014) (holding that a district court’s award of attorneys’ fees will no longer be reviewed *de novo*, but rather will now be reviewed on appeal under the more deferential “abuse of discretion” standard).

Comment

This Principle is meant to make two points clear. First, the formulation that a claim “is improper *at whatever point in time*” it becomes clear that it lacks merit is meant to emphasize the current law’s recognition that the duty to assess the viability of a claim or defense in good faith is ongoing. Claims or defenses that were appropriate at the outset may turn out to be meritless and should not be pursued once that becomes clear. This is not to say that any adverse ruling at the trial level makes pursuit or defense of a claim improper, as a good faith basis for appeal may often exist, and maintenance of the case or its defense is proper under such circumstances. However, an adverse ruling is an occasion on which the continuation of the lawsuit or its defense should be assessed, and where no good faith basis for appeal exists, the lawsuit or its defense should not be maintained.

Second, the aspect of the guideline that deals with addressing the circumstances under which initiating or maintaining a claim or defense becomes improper – namely, when competent counsel (or a party advised by competent counsel) would recognize the lack of merit – is meant to hold parties to a standard of conduct consistent with professional norms, and to eliminate “white heart, empty head” as an excuse for any otherwise inappropriate claims or defenses.

Finally, it must be emphasized that the mere fact that a claim or defense is abandoned by a litigant is not a basis for inferring that the claim or defense was improperly asserted or maintained prior to that time. Claims or defenses may be abandoned during a case for many perfectly legitimate reasons unrelated to their merits. Even where a claim or defense is abandoned because it is lacking in merit, this is, as a general matter, behavior that is to be encouraged rather than punished. As with everything else, there may well be exceptions to this generalization. For example, where a claim or defense is improperly pursued long past the point when its lack of objective validity had become clear, simply for its *in terrorem* value, such conduct may be sanctionable. The point is simply that the mere abandonment of a claim or defense is not in and of itself evidence of anything improper.

Principle V-6: Indicators that a reasonable person in the position of a party and/or the party’s counsel would know that the initiation or maintenance of a claim or defense is improper include, but are not limited to, the following:

1. **The claim or defense rests on a construction of a claim limitation that (a) was explicitly disclaimed during prosecution or in the specification, or (b) is objectively inconsistent with the plain meaning of the limitation, and the plain meaning of the limitation is not disclaimed elsewhere in the intrinsic record;**
2. **The party or original patentee (where the original patentee is not a party) or its counsel previously had made a statement about the patent to a court, the Patent and Trademark Office, or another administrative body that cannot reasonably be reconciled with the initiation or the maintenance of a claim or defense;**
3. **There is evidence (a) which establishes as a matter of law that a claim or defense is objectively baseless, and (b) which, after the initiation of a lawsuit, is actually called to the party’s attention**

through discovery, or, prior to the initiation of a lawsuit, was obtained, or through the exercise of reasonable diligence could have been obtained, from the public record or from witnesses under the control of the party; or

- 4. There is a reasonable basis to believe that a case was brought for the purpose of obtaining a settlement of a meritless claim for materially less than the likely cost of litigation.**

Comment

The listing of “indicia” of improper conduct is intended to provide guidance as to particular practices which should be viewed with skepticism. Broadly speaking, indicators 1 and 2 are meant to suggest that a court should view skeptically a litigant who seeks to turn a blind eye to the clearly formulated public record concerning the scope of the patent claims being asserted. Indicator 3 is meant to suggest that a court should view skeptically a litigant who seeks to turn a blind eye to obviously inconvenient facts. Indicator 4 is meant to suggest that a court should view skeptically a litigant who gives the appearance of abusing the litigation process.

The indicators, however, are not meant to be dispositive of the question of when a claim or defense is improperly made or maintained, as there may be facts that justify or explain any *prima facie* impropriety. Thus, for example, with respect to indicators 1 and 2, there are clearly cases where a good faith dispute exists as to what has been “objectively” disclaimed by the language of the limitation in question, the specification, the prosecution, or by statements in other proceedings.

Likewise, with respect to indicator 3, undisputed facts that are dispositive on one claim construction may not be dispositive on a different claim construction; and even where judgment is properly entered as a matter of law, the question of whether a court should do so may be a close one. Further, a litigant does not act in a sanctionable manner when it fails to recognize the significance of facts that are buried in a massive discovery record and not disclosed in expert reports, contention interrogatories, or other pleadings designed to set forth a party’s position on the merits.

Finally, the mere fact that a litigant negotiates multiple cost of litigation settlements may simply be indicative of a valid and infringed patent of limited economic significance or of a calculated funding strategy to support a good faith claim against one or more particularly significant infringers.

In short, we characterize the listed fact patterns as *indicators* both because they describe circumstances that ought to cause a presiding judge to inquire as to the bona fides of a claim or defense, and because they do not necessarily dictate what conclusion that inquiry ought to reach.

VI. CONCLUSION

In conclusion, the principles and best practices recommended by this paper represent Working Group 9’s guidance with respect to important patent damages issues. This paper obviously does not attempt to address all patent damages issues, or even to provide comprehensive coverage for the issues it does address. Instead, the paper addresses

currently debated issues for which Working Group 9 feels well-situated to propose solutions that would move the law forward, albeit incrementally, in a reasoned and just way.

The most significant departure from the current reasonable royalty framework is the new Retrospective Model of reasonable royalty damages, which: (1) reduces the uncertainty resulting from an *ad hoc* application of the “book of wisdom”; and (2) better captures the statutory requirement that impliedly suggests a retrospective look at the circumstances of the infringement and the resulting damage – i.e., that damages be “adequate to compensate for infringement.”

The paper also provides practical guidance for those involved in the nuts and bolts of patent litigation. For example, the paper sets forth a series of best practices regarding early disclosure of damages contentions, early resolution of the challenges to damages theories, and proposes a rubric for evaluating the comparability of patent licenses. Finally, there are a series of proposals for trial and posttrial considerations – relating to bifurcation, to trial time, to posttrial royalties, and lastly, to the question of awarding attorneys’ fees in appropriate cases – because these issues are important to the just resolution of these complex patent damages disputes.

THE SEDONA CONFERENCE COMMENTARY ON INFORMATION GOVERNANCE*

*A Project of The Sedona Conference
Working Group on Electronic Document
Retention & Production (WG1)*

Author:

The Sedona Conference

Editor-in-Chief

Conor R. Crowley

Drafting Team

Keith M. Angle
Jason R. Baron
Christopher Beahn
Bennett B. Borden
Howard Feldman
Liam A. Ferguson

Dean Gonsowski
Jack Halprin
Tim Hart
Virginia H. Johnson
Wayne C. Matus
Tim Noonan
Cheryl Pederson

Charles R. Ragan
Jim Shook
Peter Sloan
David L. Stanton
Cheryl Strom
Jeane A. Thomas

Thanks go to all who participated in the dialogue that led to this Commentary.

We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

PREFACE

Welcome to The Sedona Conference *Commentary on Information Governance*, a project of The Sedona Conference Working Group One on Electronic Document Retention & Production (WG1). WG1 is best known for its ground-breaking publication, *The Sedona Principles Addressing Electronic Document Production*, and as such, is generally associated in the minds of legal professionals and the public at large with civil litigation, and more specifically, with electronic discovery. But when *The Sedona Principles* were being drafted ten years ago, members of WG1 immediately recognized that no discussion of electronic discovery in civil litigation was complete, or even possible, without a discussion of the records and information management context from which requests for and responses to electronic discovery emanate. As a consequence, *The Sedona Principles* have been augmented over the past decade by WG1 commentaries that discuss the management of electronic information in the day-to-day conduct of business, government, and private life. These commentaries have included:

- *The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age*
- *The Sedona Conference Commentary on Email Management*
- *The Sedona Conference Commentary on Inactive Information Sources*
- *The Sedona Conference Primer on Social Media*
- *The Sedona Conference Best Practices Commentary on Search & Retrieval Methods*
- *The Sedona Conference Commentary on Finding the Hidden ROI in Information Assets*

With the exception of the final title in the above list, one could still sense in all these commentaries that the litigation risk management tail might be wagging the information management dog. The final Commentary on *Finding the Hidden ROI in Information Assets* broke cleanly with that history, initiating a discussion that went beyond managing the e-discovery risks associated with information, to better leverage the enormous value of information that is caught up within firms and organizations of all types.

We now take the next step, and that is to define Information Governance as an organization's coordinated, interdisciplinary approach to satisfying information compliance requirements and managing information risks while optimizing information value. In drafting this Commentary, it has been the mission of WG1 to bring together lawyers, records and information managers, technical experts, privacy and security professionals, business process engineers, human resource officers, and others, to develop a comprehensive set of basic principles to guide the development and operation of a robust Information Governance program in any organization.

The Commentary represents the collective efforts of many individual contributors. On behalf of The Sedona Conference, I wish to thank everyone involved in devoting their time and attention during the drafting and editing process, and in particular Keith Angle, Jason Baron, Dean Gonsowski, Tim Hart, Wayne Matus, Cheryl Pederson, Chuck Ragan, Jim Shook, Peter Sloan, David Stanton, and Cheryl Strom. I especially acknowledge the tireless evangelism of Editor-in-Chief Conor R. Crowley, who not only spent countless hours on the draft of this Commentary but also patiently explaining the concept of Information Governance to sometimes resistant stakeholders, helping them break out of their professional "silos" and recognize the need for a broader vision.

The Commentary represents the collective wisdom of a score of highly-qualified Information Governance professionals who contributed to the draft. The members of The Sedona Conference Working Group Series were able to review and comment on this Commentary prior to publication, it was presented at the 2013 Georgetown Law Center eDiscovery Institute, and it benefited from a six-month public comment period. But Information Governance is still very much an evolving concept. The drafters and contributors all agree that through shared experience and dialogue, Information Governance will mature as a discipline, necessitating a second edition of this Commentary. You are invited to join the dialogue online at <https://thesedonaconference.org> or submit comments by email to info@sedonaconference.org.

Kenneth J. Withers
Deputy Executive Director
The Sedona Conference
October 2014

TABLE OF CONTENTS

Principles of Information Governance (Summary)	129
Executive Summary.....	130
The Information Governance Imperative	131
Principles of Information Governance (Commentary)	137
Appendix A: Intersections.....	156
Appendix B: Maturity Continuum as it Relates to Independence.....	160
Appendix C: Risks Associated with Digital Assets.....	163
Appendix D: The Quantitative/ROI Business Case.....	166

THE SEDONA CONFERENCE
PRINCIPLES OF INFORMATION GOVERNANCE

1. Organizations should consider implementing an Information Governance program to make coordinated decisions about information for the benefit of the overall organization that address information-related requirements and manage risks while optimizing value.
2. An Information Governance program should maintain sufficient independence from any particular department or division to ensure that decisions are made for the benefit of the overall organization.
3. All information stakeholders should participate in an organization's Information Governance program.
4. The strategic objectives of an organization's Information Governance program should be based upon a comprehensive assessment of information-related practices, requirements, risks, and opportunities.
5. An Information Governance program should be established with the structure, direction, resources, and accountability to provide reasonable assurance that the program's objectives will be achieved.
6. The effective, timely, and consistent disposal of physical and electronic information that no longer needs to be retained should be a core component of any Information Governance program.
7. When information governance decisions require an organization to reconcile conflicting laws or obligations, the organization should act in good faith and give due respect to considerations such as privacy, data protection, security, records and information management, risk management, and sound business practices.
8. If an organization has acted in good faith in its attempt to reconcile conflicting laws and obligations, a court or other authority reviewing the organization's actions should do so under a standard of reasonableness according to the circumstances at the time such actions were taken.
9. An organization should consider reasonable measures to maintain the integrity and availability of long-term information assets throughout their intended useful life.
10. An organization should consider leveraging the power of new technologies in its Information Governance program.
11. An organization should periodically review and update its Information Governance program to ensure that it continues to meet the organization's needs as they evolve.

EXECUTIVE SUMMARY

Information is crucial to modern businesses. Information can have great value, but also pose great risk, and its governance should not be an incidental consideration. Despite these realities, there is no generally accepted framework, template, or methodology to help organizations make decisions about information for the benefit of the organization rather than any individual department or function.

“Information Governance” as used in this Commentary means an organization’s coordinated, inter-disciplinary approach to satisfying information compliance requirements and managing information risks while optimizing information value. As such, Information Governance encompasses and reconciles the various legal and compliance requirements and risks addressed by different information-focused disciplines, such as records and information management (“RIM”),¹ data privacy,² information security,³ and e-discovery.⁴ Understanding the objectives of these disciplines allows functional overlap to be leveraged (if synergistic); coordinated (if operating in parallel); or reconciled (if in conflict).⁵

The position of The Sedona Conference is that Information Governance should involve a top-down, overarching framework, informed by the information requirements of all information stakeholders that enable an organization to make decisions about information for the good of the overall organization and consistent with senior management’s strategic directions.

This paper explains the need for a comprehensive approach to Information Governance. The paper addresses:

- Why traditional, siloed approaches to managing information have prevented adequate consideration of information value, risk, and compliance for the organization as a whole;

1 **Records and Information Management** is the standardized process to create, distribute, use, maintain and dispose of records and information, regardless of media, format or storage location, in a manner consistent with an organization’s business priorities and applicable legal and regulatory requirements. RIM principles also provide for the temporary suspension of policies or processes that might result in the deletion of records or information subject to a legal hold.

2 **Data Privacy** is the right to control the collection, sharing and destruction of information that can be traced to an individual. In general, data privacy is more comprehensively protected outside of the United States, particularly in the European Union member states, where the Data Protection Directive provides significant restrictions on the processing and transfer of personal data, and other countries including Argentina, Canada, Israel, Switzerland and Uruguay. *See* Directive 95/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L 281/31. In the US, the approach to data privacy is generally contractual, and does not enjoy the same level of generic legal protections. Disparate laws in the United States do, however, mandate protections for specific types of data or target different groups. Examples include: patient records under the Health Insurance Portability and Accountability Act (“HIPAA”), financial information under the Graham-Leach-Bliley Act (“GLBA”), and prohibitions on the collection of information about children younger than 13 years old, under the Children’s Online Privacy Protection Act (“COPPA”).

3 **Information Security** is the process of protecting the confidentiality, integrity, and availability of information and assets, enabling only an approved level of access by authorized persons, and properly disposing of such information and assets when required or when eligible. Information security often focuses on limiting access to certain types of information that is important to the organization by restricting access through various controls including physical safeguards, technical access controls (e.g., permissions to Read, Write, Modify, Delete, Browse, Add, and Rename), authorization challenges (e.g., usernames and passwords) and encryption technologies. Security requirements can be mandated by law (e.g., HIPAA Security Rule), by contract, by industry requirements (e.g., PCI) or simply by company requirements and best practices.

4 **Electronic Discovery** (“e-discovery”) is the process of identifying, preserving, collecting, preparing, analyzing, reviewing, and producing electronically stored information (“ESI”) relevant to pending or anticipated litigation, or requested in government inquiries. E-discovery includes gathering ESI from numerous sources, reviewing and analyzing its relevance and the applicability of any privileges or protections from disclosure, and then producing it to an outside party.

- How hard costs, soft costs, opportunity costs, and risk accumulate for organizations lacking adequate control of information;
- The definition of Information Governance, its fundamental elements, and the resulting benefits to the organization; and
- The crucial role of executive sponsorship and ongoing commitment.

THE INFORMATION GOVERNANCE IMPERATIVE

We live and work in an information age that is continually – and inexorably – transforming how we communicate and conduct business. Regardless of an individual organization's size, mission, marketplace or industry, information is a crucial asset for all organizations; and if inadequately controlled, a dangerous source of risk and liability.

Some examples illustrate the highly public repercussions of information control lapses:

- Significant and increasing costs of complying with e-discovery obligations;
- Data privacy and security breaches, such as a global electronics company attributing \$171 million in out-of-pocket remediation costs to a data breach affecting 100 million persons, with the total harm, including reputational injury, estimated to exceed \$1 billion;⁶
- E-discovery sanctions, such as an award of \$8.5 million in monetary sanctions against patent holder for willfully failing to produce tens of thousands of discoverable documents;⁷
- Recordkeeping compliance penalties, such as a national clothing retailer fined over \$1 million by the U.S. Immigration and Customs Enforcement Agency for information compliance deficiencies in its I-9 employment verification system, and a retail pharmacy chain reaching an \$11 million settlement with the U.S. Government for record-keeping violations under the Controlled Substances Act.⁸

Behind the headlines, however, is a more pervasive problem – the commonly unmeasured aggregation of hard costs, soft costs, opportunity costs, and risk borne by organizations that fail to effectively control their information.

⁵ See Appendix A for additional discussion of the intersections of these disciplines.

⁶ Mathew J. Schwartz, *Sony Data Breach Cleanup to Cost \$171 Million*, INFORMATIONWEEK SECURITY, May 23, 2011, <http://www.informationweek.com/security/attacks/sony-data-breach-cleanup-to-cost-171-mil/229625379>.

⁷ *Qualcomm, Inc. v. Broadcom Corp.*, No. 05cv1958-B (BLM), 2008 WL 66932 (N.D. Cal. January 7, 2008) *vacated in part by Qualcomm v. Broadcom Corp.*, No. 05CV1958-RMB (BLM), 2008 WL 638108 (N.D. Cal. March 5, 2008); see also *Day v. LSI Corp.*, No. CIV 11-186-TUC-CKJ, 2012 WL 6674434 (D. Ariz. Dec. 20, 2012) (awarding partial default judgment and attorney's fee award of \$10,000, resulting from the loss of information that should have been retained according to both a document retention policy and a litigation hold that was not properly enforced); *Pillay v. Millard Refrigerated Servs., Inc.*, No. 09 C 5725, 2013 WL 2251727 (N.D. Ill. May 22, 2013) (issuing adverse inference instruction against a company for failing to stop the automatic deletion of employee productivity tracking data, which it had used as a reason for terminating a disabled employee).

⁸ Immigration and Customs Enforcement, Department of Homeland Security, *Abercrombie and Fitch Fined after I-9 Audit*, (2010), <http://www.ice.gov/news/releases/1009/100928detroit.htm> (last visited Nov. 13, 2013); Debbie Cai, *DOJ: CVS to Pay \$11 Million to Settle Claims of Bad Record-Keeping*, THE WALL STREET JOURNAL, (April 3, 2013), available at <http://online.wsj.com/article/BT-CO-20130403-710237.html>.

Knowingly or not, organizations face a fundamental choice: they can control their information, or by default, they can allow their information to control them.

Siloed Approaches Fail to Govern Information

Many organizations have traditionally used siloed approaches when managing information, resulting in decisions being made without sufficient consideration of information value, risk, or compliance for the organization as a whole. Examples of these silos include the various departments or administrative functions within the organization that deal with the organization's information, such as IT, Legal, Compliance, Records and Information Management, HR, Finance, and the organization's various business units. Each business unit or administrative function commonly has its own information governance policies and procedures, as well as disparate data systems and applications.

Another type of information silo consists of those disciplines that deal with specialized categories of information issues, such as data privacy and security (focused on protection of regulated classes of information), litigation e-discovery (focused on preservation and production of information in litigation), and data governance⁹ (focused on information reliability and efficiency). Over time, these disciplines have developed their own terminologies and frameworks for identifying issues and addressing specific information challenges.

The core shortcoming of the siloed approach to governing information is that those within particular silos are constrained by the culture, knowledge, and short-term goals of their business unit, administrative function, or discipline. They perceive information-related issues from the vantage point of what is familiar and important specifically to them. They often have no knowledge of gaps and overlaps in technology or information in relation to other silos within the organization. There is no overall governance or coordination for managing information as an asset, and there is no roadmap for the current and future use of information technology.

Siloed decisions concerning information often have unintended consequences for the organization as a whole, with significant cost and risk repercussions:

- An organization's individual business units independently make decisions about implementing information technology tools and systems, separate from the other business units. This results in duplication of technology and unneeded expense, and also prevents the efficient sharing of information, a valuable asset, across the organization.
- The IT Department establishes email account volume limits to relieve operational stress on an organization's email system. This results in personnel

9 We recognize that various definitions of "information governance" have been advanced (*see e.g.*, Charles R. Ragan, *Information Governance: It's a Duty and It's Smart Business*, 19 RICH. J.L. & TECH. 12 at 30-33 (2013), available at <http://jolt.richmond.edu/v19i4/article12.pdf>, and that there is an emerging discipline called "data governance," and submit that data governance is a subset of our information governance concept. The Data Governance Institute, self-described as a mission-based and vendor neutral authority on essential practices for data strategy and governance, defines "data governance" as "a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods." *Definitions of Data Governance*, THE DATA GOVERNANCE INSTITUTE, http://www.datagovernance.com/adg_data_governance_definition.html (last visited Nov. 13, 2013). So viewed, "data governance" does not address "why" an organization chooses to do certain things with its data and other information; that is the critical role of Information Governance, ensuring that actions users take with information-related assets is consistent with organizational strategy.

moving email to storage on local drives and devices, exacerbating both data security risks and difficulties in finding and preserving such email for litigation.

- Legal counsel issues overbroad litigation holds to avoid even a remote possibility of spoliation sanctions. This results in excessive costs in pending and future litigation and also the unnecessary retention of data.
- Personnel are allowed to conduct an organization's business on their own laptops and smartphones, under a Bring-Your-Own-Device ("BYOD") program to increase convenience and efficiency but without sufficient BYOD policies and controls or planning for natural attendant consequences. This results in data security exposures and difficulties in applying records retention policies and in preserving and collecting data for litigation.
- Privacy and data security controls are applied to an organization's service providers, but are not used to ensure that service providers also meet the organization's records retention requirements. This may result in inconsistent application of such requirements to records.
- Records manager initiates a robust data and email retention program without regard to potential technological limitations or the burden associated with retaining, searching and reviewing the resulting data for e-discovery purposes.

In the post-Sarbanes-Oxley world, many companies have adopted codes of conduct, in which they broadly proclaim that the organization and its employees comply with all applicable laws (including privacy and data security requirements), protect confidential information, use electronic communications wisely, and follow procedures for retaining records. The siloed approach to addressing information issues, however, inevitably spawns a multitude of information-related policies adopted through various projects and initiatives. Thus, rather than a clear, uniform set of information policy guidance, employees face a cacophony of conflicting policies and procedures, making compliance virtually impossible in the heat of a competitive business environment, and negatively impacting productivity.

The "elephant in the room" is the organization's need to harness and control its information, coupled with the inadequacy of a siloed approach for accomplishing this crucial goal. The solution to this quandary is for organizations to find a way to bridge across their silos, so that issues of information compliance, risk, and value can be identified, understood, and addressed for the benefit of the entire organization.

Information Governance

"Information Governance" as used in this Commentary means an organization's coordinated, inter-disciplinary approach to satisfying information legal and compliance requirements and managing information risks while optimizing information value. Organizations that adopt Information Governance programs are able to bridge across silos, thereby perceiving and understanding information-related issues from the perspective of the overall organization. Information Governance also helps ensure that decisions and solutions regarding information compliance, risk controls, and value optimization will serve the needs of the entire organization rather than the insular needs of individual silos.

To accomplish Information Governance, organizations should:

- Establish a structure for Information Governance, which will vary in form depending on the organization's size, complexity, culture, and industry and regulatory environment;
- Determine the organization's strategic objectives for Information Governance, based upon a comprehensive assessment of information-related practices, requirements, risks, and opportunities;
- Reconcile the various compliance requirements and risks addressed by different information-focused disciplines, such as records and information management, privacy, data security, and e-discovery; and
- Implement an Information Governance program with the structure, direction, resources, and accountability to provide reasonable assurance that the program's strategic objectives will be achieved.

The Benefits of Information Governance are Significant

The advantages of establishing an Information Governance program are many and varied, depending upon the information-related issues and risks an organization faces. Beyond addressing the risks above, an enterprise-wide Information Governance program will help organizations achieve the following advantages, all of which add to the bottom line:

- Business performance improvements, as users gain confidence that they can locate valuable information efficiently and reliably, and better understand how to address information-related risks;
- Realization of "option value" as the organization leverages existing information and technologies across diverse business units, consolidates technologies and administrative staff, and reduces license fees;
- More reliable and efficient processes and procedures for e-discovery;
- Reduced storage costs and administrative burdens, as obsolete and worthless information is eliminated; and
- Reduced costs and enhanced compliance with legal obligations for records retention, privacy and data security, and e-discovery, as information policies and processes are rationalized, integrated, and aligned in accord with the organization's information governance strategy.

Senior Leadership Support is Essential

The commitment of senior leadership is crucial for organizations to be successful in adopting Information Governance. Such ongoing commitment is particularly important given the challenge of effectively bridging across existing organizational silos.

Thus, senior leadership should sponsor and firmly support the organization's Information Governance efforts by:

- Endorsing the importance of Information Governance to the entire organization;
- Chartering a structure of responsibility and accountability for implementing an Information Governance program;
- Adopting or approving the strategic objectives of the Information Governance program;
- Providing appropriate resources to implement and sustain the Information Governance program;
- Establishing a supportive “tone at the top” and an environment in which Information Governance remains an organizational priority; and
- Ensuring that the Information Governance program is administered consistent with its objectives and is periodically reviewed and updated.

There is often a balance of value against cost or risk that changes over time for a given information asset. Organizations may leverage information effectively over the short term, but once the data's short-term use is expended, the data is often stored away and rarely reassessed for any long-term strategic value. Left ungoverned, this potentially valuable asset is not only wasted, it also may become a significant liability. Through proper information governance, organizations can realize additional benefit from their information assets over time while reducing risk.

The Business Case for Information Governance

Multiple business cases can be established for pursuing Information Governance. Successful adoption of the information governance approach requires both strategic commitment (adoption of information governance as an organizational priority) and also tactical efforts (such as specific projects to establish and implement the program). A business case will be needed, both to support the strategic commitment and also to justify the expenditures of time, effort, and funding required for specific implementation projects. Because the business case for information governance must be persuasive at both strategic and tactical levels, the business case should include both strategic (qualitative) and project-based (quantitative, ROI) elements.

The Strategic/Qualitative Business Case:

Information governance is an ongoing program that evolves over time through maturity levels. As such, it is unrealistic to attempt to comprehensively quantify all of its benefits. One might just as easily attempt to exhaustively measure all benefits of managing the organization's tangible or people assets. ROI analysis is best used for applications of information governance to specific, issues or projects within the information governance initiative, as discussed in Appendix D.

At a strategic level, the business case should instead convey how information governance aligns with and amplifies the core values and fundamental, strategic objectives of the organization. For example:

- **Low Cost Provider**

Companies singularly focused on operational efficiency and cost control, such as in low-margin, high-volume industries or market segments, may adopt information governance to streamline information workflows and reduce unnecessary information storage and retention, thereby reducing costs and increasing business efficiency.

- **Innovative Excellence**

Organizations driven by creative innovation and excellence in products and services may adopt information governance to maximize the value of their information assets, helping them capture valuable information for innovative repurpose while minimizing the distraction of unnecessary information.

- **Trusted Provider/Advisor**

Organizations with the core value and brand of being a trusted business provider or advisor may adopt information governance to strengthen their protection of information that customers or clients entrust to the organization and also to enhance third-party perceptions of the organization as a trusted custodian for such information.

- **Integrity/Ethics**

Companies, including publicly traded organizations and those in highly-regulated industries, may adopt information governance as a complement to their internal control systems and corporate ethics and integrity programs to ensure information-related legal compliance and risk management.

In each of the above examples, information governance provides specific, tangible benefits that often can be quantified on an ROI basis as discussed below. Yet, in each example, information governance also amplifies the organization's core value of choice, by ensuring that information is handled in alignment with the strategic value or brand. This alignment allows information governance to reinforce the particular organization's fundamental values, as information is managed in a way that "walks the walk."

Conversely, information governance also helps organizations avoid cultural dissonance for their core values, such as, for example, the "low cost provider" that squanders money on information inefficiency and unnecessary retention; the "innovative excellence" company that fails to optimize the value of its information; the "trusted partner/provider" that is careless with the information entrusted to it; or the company espousing "integrity and ethics" that fails to establish a control environment for information as a valuable asset and as a means to detect and prevent compliance lapses. Thus, adoption of information governance can have profound, strategic significance beyond the quantitative ROI measures mentioned below and considered in more detail in Appendix D.

The Quantitative/ROI Business Case:

A typical ROI analysis weighs the benefits of a particular project against its cost, and calculates the length of time it will take to recoup the cost. The quantitative aspects of the business case are best determined by focusing on specific applications of information governance to identified problems or opportunities, or to discrete projects for implementation of the Information Governance program.¹⁰

The quantifiable benefits from pursuing information governance generally fall into four main categories: optimizing corporate value, risk reduction, hard cost avoidance, and soft cost avoidance. See Appendix D for factors to consider when building a quantitative business case with these ROI categories.

THE SEDONA CONFERENCE
PRINCIPLES OF INFORMATION GOVERNANCE

Principle 1. Organizations should consider implementing an Information Governance program to make coordinated decisions about information for the benefit of the overall organization that address information-related requirements and manage risks while optimizing value.

Organizations benefit in several ways from managing information as a valuable asset. In order to realize these benefits, an Information Governance program should be established in a manner consistent with the organization's industry, compliance, and risk environments.

Any Information Governance program should incorporate the following principles: transparency, efficiency, integrity, accountability, and compliance. To be successful, the Information Governance program must be sponsored and firmly supported by the organization's senior leadership.

A core component of any Information Governance program should include a comprehensive data classification capability, combined with the effective, timely deletion of information. By taking a comprehensive approach to identifying and addressing information-related requirements, organizations can ensure compliance needs are met and conflicting issues are considered. It is also helpful to identify and assess information risks, such as user access control (information security) and system failure (business continuity and disaster recovery), and ensure that such risks are understood so effective information controls can be put in place. This approach also aids in understanding information-related strategic and operational objectives to help ensure that information value can be optimized without compliance lapses or uncontrolled risk.

Although there are many stakeholders with divergent interests in managing information, decisions about governing information should benefit the overall organization, rather than a particular department or discipline.

To enable an organization to make coordinated decisions about information for the benefit of the organization, the primary responsibility of an Information Governance

¹⁰ See generally, S. Soares, *Selling Information Governance to the Business: Best Practices by Industry and Job Function* (2011) (providing insight into the best ways to encourage businesses to implement an information governance program).

program should be to create and maintain processes and procedures necessary for a coordinated, overall approach to decisions about information. If agreement cannot be reached among stakeholders, the Information Governance program should provide a method for decisions to be made (subject to a challenge process) to enable the organization to move forward. Transparency, efficiency, integrity, accountability, and compliance are integral to the ability to perform this overall coordination and tie-breaking function successfully.

Responsible decision makers should use the Information Governance program at the time they make decisions about information. Care should be taken to design the Information Governance program so that it can be used in this way. Existing governance mechanisms (such as budgetary governance or systems approval) may not be designed for users to interface with at the time decisions are being made. However, these can be leveraged or modified or new ones may be created, depending on an organization's circumstances.

Principle 2. An Information Governance program should maintain sufficient independence from any particular department or division to ensure that decisions are made for the benefit of the overall organization.

The information governance function must focus on the best interests of the organization. In order to fairly and effectively balance needs, however, the information governance program should have meaningful and balanced input from such departments as IT, legal, compliance, RIM, and the business units. One approach to accomplish this is to designate an executive who has sufficient independence to balance the competing needs of stakeholders rather than the interests of a single department. Ideally, the executive in charge of the Information Governance program reports at the same level as a General Counsel, CCO, CFO, or CIO. Another way to make decisions for the benefit of the overall organization is through a committee that has representation from impacted stakeholders, coupled with a process for elevating disagreements to a chief executive. Such a structure should be the ultimate goal for organizations with mature Information Governance programs. However, many organizations do not currently have in place any overarching information governance structure and their initial steps may include assigning information governance responsibilities to designated individuals within departments or lines of business. As this is not the optimal governance structure to reap the benefits of a coordinated approach to information governance, organizations should strive for a structure that results in meaningful and balanced input from all impacted departments or divisions as their Information Governance programs mature.¹¹

Many organizations have various departments (i.e., business units, IT, Legal, etc.) that take direction from a CEO or COO. Because goals differ across departments or functions, conflicts of interest may arise if the executive responsible for the Information Governance program reports to an individual stakeholder department.

An Information Governance program should ensure that decisions about information are made in the organization's best interests. Deciding for the overall good of the organization involves balancing the sometimes competing interests of many stakeholders. This balancing creates the potential that a given decision may not align with the particular objectives of a given department, particularly when the decision involves a

¹¹ See Appendix B for a discussion of the Information Governance Maturity Continuum.

balancing of cost and risk. For example, one stakeholder may believe a cloud-hosted service will reduce the cost of storing information, but another may perceive an increased risk associated with the data being hosted in the cloud. The reduced cost may be attractive to a department such as IT, and the increased risk may be unattractive to another department such as Legal. In many cases, stakeholders can arrive at a mutually agreeable position that maximizes the benefit to the overall organization, for example by implementing mitigation steps that decrease the risk to one department without substantially increasing the cost to other departments.

Though it is appropriate for departments to operate autonomously in carrying out their primary function, decisions about information governance should be coordinated across all departments and stakeholders as they impact the organization as a whole. Because such decisions require an overall balancing between the needs and interests of different stakeholders, it is important for the information governance function to be independent within the organization.¹²

Principle 3. All information stakeholders should participate in an organization's Information Governance program.

Information Governance programs should seek to be inclusive and to involve all parts of an organization (business units, departments, etc.) that have an interest in the company's information.¹³ This may require involvement from all of the organization's departments or business units, which may require different levels and types of activity from stakeholders.

An inclusive process will ensure that decisions about information represent all viewpoints, identifying and resolving potential conflicts early and prior to any action being taken that could have an adverse impact to the organization. For example, an organization might consider a policy that bans MP3 (audio) files from being stored on company resources because they are often identified as unauthorized employee music collections, but there may be cases where such files contain training webcasts and may be needed by HR or corporate training. Without involvement of all parties, valuable information could be lost and adversely impact the organization.¹⁴

However, participation does not require a "seat at the table" for every person or even every department with an interest in the organization's information. In larger organizations, active participation from every group could create an unwieldy team unable to reach decisions. A more effective approach would be to design an appropriate structure or methodology to ensure that all stakeholder interests are represented. An organization could create a process to identify groups with common interests, appoint certain committee members as proxies for other groups, or design surveys or feedback sessions to ensure that all interests are adequately identified and represented.

¹² For further explanation, see Appendix B.

¹³ Cf. The Sedona Conference, *Finding the Hidden ROI in Information Assets*, February 2011, <https://thesedonaconference.org/download-pub/466>.

¹⁴ *Equal Employment Opportunity Commission v. Ventura Corp. LTD.*, Civ. No. 11-1700, 2013 WL 550550 (D.P.R. Feb. 12, 2013) (finding that even though there was no evidence of bad faith, a company that failed to preserve pertinent emails and hiring-related documents when it migrated to a new software system and restructured its office, ignored repeated requests to preserve the documents, and retained relevant emails that highlighted its missteps in preserving evidence amounted to spoliation that permitted sanction, exclusion of evidence, and an adverse inference instruction).

In most organizations, stakeholders from the core disciplines of records and information management, data privacy, information security, data governance and e-discovery should be represented in the Information Governance program. These disciplines will involve IT, Legal/Compliance, Risk, Audit and RIM functions. Representatives of lines of business and core operational functions should also be included to ensure that the practical needs of the organization are properly considered. It is important to include core operational functions that have unique information governance issues. For example, human resources and environmental functions typically have legally mandated retention for some of their information.

Principle 4. The strategic objectives of an organization's Information Governance program should be based upon a comprehensive assessment of information-related practices, requirements, risks, and opportunities.

An effective Information Governance program should be designed, implemented, and monitored based upon organization-wide objectives established from a comprehensive assessment of the interests and concerns of key stakeholders within the organization, such as IT, Legal, Compliance, Records and Information Management, and various business units. The program objectives should address and coordinate the stakeholders' existing practices and approaches to issues such as records and information management, privacy and data security, and litigation preservation; and reconcile the practices and approaches with applicable legal requirements. Other major responsibilities of the Information Governance program should include gathering stakeholder requirements, such as those needed to create and publish requirements. Although the Information Governance program does not own the requirements, it owns responsibility for collecting requirements and considering them to arrive at a decision for the good of the organization overall.

To determine its information-related practices, requirements, risks, and opportunities, an organization should first identify the various types of information in its possession, custody or control, assess whether it owns the information or possesses it for third-parties; and determine whether the information is held by the organization, by third-parties for the organization, or both. The organization should next identify its current information lifecycle practices, including practices pertaining to:

- Creation and/or receipt of information;
- Determining location and media for storing information, including in both active and inactive environments;
- Disaster recovery and business continuity;
- Security for private or confidential information;
- Retention of information in both active and inactive environments;
- Implementation, maintenance and release of legal holds due to litigation or government proceedings; and
- Disposal/destruction of information.

A review of existing written policies, procedures, retention schedules, data maps and contractual arrangements is helpful in identifying and understanding these information-related practices. However, input from the organization's information stakeholders, including IT, Legal, Compliance, Records and Information Management, and business units, among others, is also essential to gaining an accurate and complete understanding of both the strengths of current information governance practices and areas where improvement may be necessary.

Organizations can then assess their identified information types and related practices in light of information opportunities, risks and compliance requirements including:

Opportunities

- Reducing costs and risks of complying with e-discovery obligations, by decreasing the volume of unnecessary information, understanding where information is stored, and considering e-discovery costs and risks when approving locations or formats for creating or storing information;
- Utilizing information to support evidence-based decision making;
- Optimizing accessibility of information to enhance productivity and efficiency;
- Realizing cost savings by decreasing the volume of unnecessary information, and rationalizing storage options to better meet demands while reducing cost;
- Enabling access to information for new and valuable combinations and uses;
- Enhancing the organization's reputation as a trusted custodian of PHI, PII, and other classes of protected information; and
- Achieving cost savings and reducing risk through efficient and appropriately-scoped preservation of information for litigation or government proceedings.

Risks

- Loss of records or other valuable information;
- Loss of integrity, authenticity, and reliability of records or other valuable information;
- Unavailability of information vital to the organization's continued operation;
- Accumulation of information (both by the organization and third parties) not (i.e., never or no longer) required for legal compliance or business needs;
- Creation or storage of information in locations or formats that increase the risk or cost of e-discovery, without a corresponding business benefit to outweigh the increased risks and costs;

- Creation of internal RIM requirements that are not followed;
- Breach of PHI, PII, or other classes of protected information;
- Harm to information from malicious access or attack;
- Inability or failure to detect and respond effectively to data breaches;
- Loss of intellectual property protection;
- Loss of privilege or confidentiality of information;
- Failure to preserve information relevant to litigation or government proceedings;
- Over-preservation of information for litigation or government proceedings; and
- Failure to release information (held by the business, by the legal department, or by outside vendors like law firms, expert witnesses, review vendors, etc.), from preservation once no longer relevant to litigation or government proceedings.

Compliance Requirements

- Legal and contractual requirements for:
 - Records creation, retention, management, and disposition;
 - Privacy and security for PHI, PII, and other classes of protected information;
 - Protection of intellectual property and confidential information; and
 - Preserving information relevant to litigation or government proceedings.

These considerations will differ between jurisdictions, industry sectors, and organizations; and among organizations, there will be a range of risk tolerances and cultures regarding these matters. Industry standards, maturity models, and benchmarking data for comparable organizations are useful considerations for this assessment.¹⁵

¹⁵ Useful standards and models include:

- International Organization for Standardization, *Information and Documentation-Management Systems for Records - Fundamentals and Vocabulary*, ISO 30300:2011 (2011).
- International Organization for Standardization, *Information and Documentation - Records Management - Parts 1 and 2*, ISO 15489-1:2001 (2001); ISO 15489-2:2001 (2001).
- International Organization for Standardization, *Information Technology - Security Techniques*, ISO/IEC 27000:2012(2012); ISO/IEC 27010:2012 (2013); ISO/IEC TR 27019:2013 (2013).
- ARMA, *Generally Accepted Recordkeeping Principles® & Information Governance Maturity Model*, <http://www.arma.org/r2/generally-accepted-br-recordkeeping-principles> (2013).
- COBIT 5, *A Business Framework for the Governance and Management of Enterprise IT* (2012), available at <http://www.isaca.org/COBIT/Pages/default.aspx>.
- The Sedona Conference, *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production* (Second Edition) (June 2007), <https://thesedonaconference.org/download-pub/81>.
- ISO standards, such as the ISO 30300 Series, *Management Systems for Records*; ISO 15489, *Records Management*; and the ISO 27000 Series, *Code of Practice for Information Security Management*.
- ARMA's *Generally Accepted Recordkeeping Principles® & Information Governance Maturity Model*.
- COBIT 5, *A Business Framework for the Governance and Management of Enterprise IT*.

An organization should use the results of the above assessment to determine its objectives for information governance. Well-framed strategic objectives for information governance can guide the design and implementation of the organization's Information Governance program, helping to clarify what elements of structure, direction, resources, and accountability will be pursued, as discussed under Principle 5. Establishing strategic objectives in this manner should clarify decision making on priorities and funding of the effort. Strategic objectives should be measurable to better ensure that progress toward them can be observed and reported. Such measures may be quantitative (i.e., data volumes or run-rates) or qualitative (i.e., assessment or audit against program standards or upon completion of transactions or litigation matters). Measurability of objectives is essential for accountability, discussed under Principle 5. Perhaps the most important feature of this exercise is that it compels organizations to look beyond the confines of traditional silos within organizations.¹⁶

Principle 5. An Information Governance program should be established with the structure, direction, resources, and accountability to provide reasonable assurance that the program's objectives will be achieved.

To provide reasonable assurance that an Information Governance program will meet an organization's strategic objectives, the program should have structure, direction, resources, and accountability. Depending on the size of the organization, responsibilities such as change management and communication to raise awareness of the information governance function, user training, creating the information governance matrix, and gathering metrics required for management control and monitoring may also be important.

Structure

One means of ensuring that an organization's various information needs are comprehensively addressed is to establish a unified framework in which the organization's various information types can be categorized according to information-related compliance requirements and risk controls. Such a framework should categorize information types by content and context.¹⁷ This will normally require input from a wide range of subject matter experts, including, for example, human resources, accounting, compliance, and environmental.

16 For example, in its information governance assessment, a financial services organization confirms that it has customer information subject to privacy and data security requirements, which it regularly transfers to the custody of various service providers in the ordinary operation of its business. From the siloed perspective of privacy and data security compliance, the organization satisfies the applicable requirements of the Federal Trade Commission's Safeguards Rule (FTC Standards for Safeguarding Customer Information, 16 C.F.R. Part 314 (2002)) by, *inter alia*, establishing internal controls for selecting and retaining service providers and by contractually requiring them to establish safeguards to ensure security for protected customer information. The organization also periodically audits its service providers to assess the effectiveness of their information security safeguards.

However, through its information governance assessment, the organization determines that its internal requirements for records retention periods are not followed by its service providers, such that some service providers retain customer information for either a shorter or longer period of time than is required under the organization's records retention schedule. The organization also determines that its legal hold process may not include certain customer information relevant to litigation that is in the custody of various service providers, yet arguably within the "control" of the organization for discovery purposes.

As a result of the assessment, the organization decides that one of its strategic objectives will be to apply information governance controls to customer information possessed by its service providers. This strategic objective will allow the organization to ensure that service providers implement appropriate safeguards to protect customer information, comply with the organization's records retention schedule and be responsive to legal holds that may be imposed upon customer information possessed by service providers.

17 Information context is significant, because different copies or instances of the same information content may be used for different purposes, thereby triggering different compliance requirements and risks. For example, a single contract may simultaneously exist in multiple instances for different purposes, including the original executed hard copy version; the scanned, digitized version that the organization declares as the official record of the contract; disaster recovery backup copies of the digitized contract; reference copies of the contract used for business convenience in various departments; and a preserved version of the contract under legal hold due to pending litigation. In each of these contexts, different compliance requirements and risks apply to the same information content of the contract.

Attached to this framework of information types are the applicable rules the organization applies to the respective information. These rules reflect legal and regulatory requirements for records retention, information management, and information security and protection. The rules reflect the organization's operational needs for how information will be retained, managed, and protected, and also the organization's risk controls. The unified framework allows the organization to identify, understand, and follow the appropriate rules for its information types.

In place of siloed structures governing data security, retention, and preservation, an organization could establish an information governance matrix. An information governance matrix is a classification structure for the organization's information types similar to a traditional records retention schedule or data security grid but which integrates all established rules governing the organization's information types. An information governance matrix is thus a repository of integrated rules for information from the organization's perspective as a whole, rather than merely one or more of its siloed functions. An information governance matrix should be designed to meet the needs of various audiences and multiple uses within the organization. It is essential, for all of the Company's business information, that the Company establish and clearly communicate responsibility for complying with the integrated rules included in this governance matrix. Otherwise, "orphan data" can greatly increase the cost and risk of e-discovery.

An organization should strive to establish a common vocabulary for its various information types.¹⁸ A common vocabulary helps ensure information is properly classified, so that the applicable rules for such information types can be identified and followed.

Direction

Organizations should communicate to all information users the organization's expectations for information governance. Vehicles commonly used by organizations to provide such direction include policies, contracts, retention schedules or information governance matrices, procedures and protocols, and guidance and training.

Many organizations have an array of policies that directly or indirectly address information governance topics. Examples include a records-and-information management policy, a communications policy, a computer use policy, an Internet and social media policy, a bring-your-own-device policy, an information security policy, and a legal hold policy. In many organizations, such information-related policies accrete over time, each designed to meet the needs of discrete stakeholders and silos of the organization. They commonly address only limited aspects of information governance and may be in conflict with each other. Organizations should identify all such existing policies, review them for inconsistencies and gaps in coverage, and reconcile them or integrate the majority of these policies into a single information governance policy. Similar to the information governance

¹⁸ Whether an organization relies upon traditional structures such as records retention schedules and data security grids or integrates them into an information governance matrix, such structures are commonly organized as taxonomies. A taxonomy is a defined hierarchy with classes and sub-classes forming "trees" of classification. In a taxonomy, it is only possible to move downward into sub-classes, or upward into super classes that subsume all of the classes below. Taxonomies are flat and linear, and therefore limiting. In contrast, ontologies link classes in a non-hierarchical way, forming associations that are non-linear. Thus, the widget purchase order may be associated hierarchically with accounting recordkeeping; but at the same time, it may also be associated with documentation of contract rights and duties, and yet other business functions. Instances of the widget purchase order information may also, simultaneously, be associated with disaster recovery restoration, with information protection issues (due to where versions of the purchase order are located physically or virtually), and with applicable legal holds. The complexity of the digital environment, in which the same information content simultaneously exists in different locations and contexts, triggering different information governance rules, makes ontology a promising perspective for applying information governance to an organization's information.

matrix, an information governance policy expresses in one place all of the organization's policy-level expectations for governance of information.

Contracts with third parties are another means of providing direction for information governance. Organizations commonly allow information to be transferred to or held by third parties, such as service providers for business operations; management, legal, accounting, and technology consultants; data hosting providers; and hard-copy records storage providers. The organization's expectations for information governance should be communicated to such third parties through its contracts with them.¹⁹ For example, engagement letters with law firms should confirm the firm's obligations to protect and preserve information, and also the company's right to require destruction or return of information after the matter or engagement is concluded.

Organizations should also have specific procedures and protocols that provide explicit direction on information creation, receipt, use, dissemination, protection, retention, preservation, and ultimate disposition. Organizations should also establish effective guidance and training regarding information governance, delivered in a way that empowers individuals to make timely, compliant decisions regarding information.²⁰ Accordingly, training and guidance resources should be tailored to meet the specific needs of recipients and should provide the concrete direction the recipients need to make information-related decisions consistent with the organization's information governance expectations.

Resources

Organizations should provide the people, technology, and implementation resources needed to support their Information Governance program and accomplish the organization's strategic objectives.

People resources include staffing of the management and administrative roles supporting the Information Governance program itself, as discussed above under Principle 3. Staffing should be commensurate with the program's scope and objectives, and roles and responsibilities should be defined. Key points of contact should be identified within the organization, and those in such roles should be accessible and responsive. People resources reflect the focus and engagement of stakeholder representatives, such as from Legal, IT, Compliance, Records and Information Management, other administrative functions, and lines of business. People resources also reflect the recognition that information governance is part of everyone's job responsibilities within the organization.

Technology resources include systems and applications used for creating, using, and storing information, into which should be placed structures and controls for information governance. Technology resources also include systems and applications for managing, tracking, and reporting regarding the Information Governance program itself. Both kinds of technology should be used for the program's scope and objectives. Information governance technology resources should be procured only after requirements for such tools have been defined, consistent with the organization's strategic objectives for

19 In some regulated sectors, contractual control of information protection by such service providers is an explicit legal requirement. For example, HIPAA covered entities must contractually require their business associates to provide compliant security for electronic protected health information (ePHI) created, received, received, maintained, or transmitted on behalf of the covered entity. 45 C.F.R. § 164.314(a).

20 *Day v. LSI Corp.*, No. CIV 11-186-TUC-CKJ, 2012 WL 6674434 (D. AZ. Dec. 20, 2012) (awarding sanctions for, among other things, failing to follow own document retention policy).

information governance. Organizations should carefully consider whether the contemplated technology can fully achieve the program's desired objectives.

Implementation resources are also needed. These include project management tools and processes to be used as elements of the organization's Information Governance program.

Accountability

The effectiveness of an Information Governance program will turn upon whether the organization establishes accountability for meeting program expectations and for achieving the organization's strategic objectives for information governance. In internal control systems, this atmosphere of accountability is the "control environment."²¹ The organization's senior leadership establishes the "tone at the top" regarding strategic objectives, the importance of reaching these objectives, expected standards of conduct, and accountability. In all forms of direction, the visible commitment and support of the organization's senior leadership is crucial.²²

Management reinforces these expectations, and the related roles, responsibilities, and accountability, across the organization. The Information Governance program should clarify roles and responsibilities, both for information users and also for those managing the Information Governance program.

Information Governance program objectives should be linked to observable and measurable outcomes; and compliance audits or comparable assessments of the program should be conducted on a regular, periodic basis, followed by appropriate corrective actions as needed. Program outcomes should be periodically compared to program objectives, and such outcomes should be tracked by those responsible for the Information Governance program.

The results of such outcome measures and program assessments should be reported periodically to the organization's senior leadership to provide reasonable assurance that the program's objectives are or will be satisfied.

Principle 6. The effective, timely, and consistent disposal of physical and electronic information that no longer needs to be retained should be a core component of any Information Governance program.

It is a sound strategic objective of a corporate organization to dispose²³ of information no longer required for compliance, legal hold purposes, or in the ordinary

21 The internal control concept of a control environment is a model that organizations may consider in pursuing information governance, particularly for establishing accountability and managing risks around specific objectives. See Committee of Sponsoring Organizations of the Treadway Commission ("COSO"), *Internal Control-Integrated Framework Executive Summary - English*, (2013), <http://www.coso.org/documents/Internal%20Control-Integrated%20Framework.pdf> ("Internal control is a process effected by an entity's board of directors, management, and other personnel, designed to provide reasonable assurance regarding the achievement of objectives relating to operations, reporting, and compliance.")

22 In some aspects of information governance, senior leadership involvement is legally required. For example, entities subject to the FTC's Red Flags Rule must obtain board-level approval of the initial Identity Theft Program, and must involve the board or senior management in the oversight, development, implementation, and administration of the Program. 16 C.F.R. § 681.1(e)(1) & (2). ISO 30300 provides that "Top management is responsible for setting an organization's direction and communicating priorities to employees and stakeholders."

23 In this Commentary, the term "disposal" will be used narrowly to refer to the final destruction or deletion of information that no longer has any regulatory, statutory, compliance, legal or operational value and is not subject to any retention or preservation requirement. The effective disposal of data should purge all copies of that information from relevant systems so that they are no longer retrievable.

course of business.²⁴ If there is no legal retention obligation, information should be disposed as soon as the cost and risk of retaining the information is outweighed by the likely business value of retaining the information. This may require a culture shift in some organizations that have developed a “keep it just in case” mentality. Typically, the business value decreases and the cost and risk increase as information ages. Timely disposal of information in a consistent and effective manner provides many benefits, including reduced storage and labor costs,²⁵ reduced costs and risks of complying with discovery obligations, and an increased ability to retrieve important organizational information. Organizations should therefore consider procedures to achieve the regular destruction of unnecessary information.²⁶ Organizations should also consider whether information considered private or confidential to third parties should be disposed of within a reasonable amount of time after it ceases to be useful to the organization in order to minimize the risk of disclosure.

While most organizations are familiar with managing paper records (and most retention schedules were drafted with paper in mind), it is important that the organization’s retention schedules account for both hard copy and electronic records. For example, record owners may find it difficult to apply the concepts original versus copies to digital information.

The term “hold” is used broadly in this commentary to cover preservation obligations that are independent from routine recordkeeping requirements, such as reasonably-anticipated or active litigation, governmental inquiries, outside audits, or contractual requirements. A hold may take the form of:

- A legal or litigation hold, i.e., the preservation of data for purposes of reasonably anticipated or active litigation or investigations;
- A tax hold, i.e., the preservation of information in ongoing audit or review of records related to tax obligations, such as financial and accounting records;
- A contractual hold is an agreed-upon obligation that an organization has with its customers, vendors, divested entities or other third parties that creates an obligation to preserve or dispose of information that exists separately from the retention schedule.²⁷

Records Retention

To create a proper data disposal process, the organization should consider all applicable legal, regulatory, and contractual requirements, in conjunction with the business

24 *Managed Care Solutions, Inc. v. Essent Healthcare*, 736 F. Supp.2d 1317, 1326 (S.D.Fla. Aug. 23, 2010) (rejecting the argument that there is no reasonable business routine demanding that data be destroyed after [13 months], especially in light of developments in the technology field (including the ability to inexpensively maintain documents at an off-site server) and industry standards stating the exact contrary.” (citing *Matya v. Dexter Corp.*, No. 97-cv-763C, 2006 WL 931870, at *11 (W.D. N.Y. Apr. 11, 2006) and *Floeter v. City of Orlando*, No. 6:05-CV-400-Orl-22KRS, 2007 WL 486633, at * 7 (M.D. Fla. Feb. 9, 2007)).

25 Though some may view data storage as a low-cost concern, the maintenance, retention and discovery-based review of unnecessary information is far from cheap. In the aggregate, storage is quite expensive. See, e.g., Jake Frazier, *‘Hoarders’: The Corporate Data Edition*, LAW TECHNOLOGY NEWS, (2012), <http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202581938140>.

26 Principle of Disposition, ARMA, *Generally Accepted Recordkeeping Principles*®, <http://www.arma.org/r2/generally-accepted-br-recordkeeping-principles> (last visited Dec.3, 2013) (“An organization shall provide secure and appropriate disposition for records and information that are no longer required to be maintained by applicable laws and the organization’s policies.”).

27 An organization should be wary of this type of obligation, as it could create onerous obligations to dispose of copies of electronic data that may not be within the control of the organization, and inconsistent obligations where different contracts prescribe different retention periods.

value of the organization's information. The organization might begin this process by evaluating its legal/regulatory requirements at all levels and across all jurisdictions relevant to its business (state, federal and/or international) and clustering those records into categories.²⁸ This exercise will enable the organization to more easily identify the appropriate retention period applicable to each category of records, while also facilitating the analysis of certain key factors relevant to the retention determination, including the cost vs. risk associated with a category of records.²⁹

It is important for the organization to remember that the operational value of a records category cannot be the sole consideration in determining a proper retention schedule; legal, regulatory and compliance objectives are of paramount concern. It is equally important, however, that operational value (e.g., maintenance of historical records, research and development processes, other business-driven objectives) be considered as the organization formulates its retention protocols. Otherwise, the organization may squander valuable opportunities to reduce cost while minimizing risk. For example, organizations should strive to avoid retaining information simply because it may possibly be useful at some point in the future and instead undertake a cost-benefit and a risk-benefit analysis with respect to each category of data it maintains, thereby ensuring that the advantages of retaining a given set of information outweigh the potential costs and risks associated with disposing of that information.

Hold/Preservation Analysis

Before the organization disposes of any business records, it should conduct a hold analysis to determine whether there are any legal/regulatory or other obligations in place that require the organization to retain information, regardless of its business value. In order to effectively identify its preservation obligations, it is advisable for the organization to develop and implement protocols designed to track legal/regulatory holds and map them to the relevant sources of information, or take other steps to label, segregate and preserve the information. A key aspect of this exercise is to communicate those protocols to the relevant individuals within the organization, and provide a point of contact (typically, a member of the legal or compliance department) who will address any questions regarding hold procedures and best practices.³⁰

It is important for the relevant constituencies within the organization – not just the legal/compliance department – to understand that a legal hold supersedes all other records and information management and retention schedules, and that a hold requires the immediate suspension of the disposal process for all affected information during the time mandated by the hold. Thus, it is critical for the organization to incorporate a “hold and release” capability into its records disposition process, so that once the hold is released or has expired, the affected information can be placed back into the appropriate retention schedule.

28 For some organizations, local, municipal and/or regional recordkeeping regulations may apply and, if so, should also be considered when developing an appropriate records retention schedule.

29 For more information, see ARMA International Standards and Best Practices, <http://www.arma.org/r2/standards-amp-best-practices> (last visited Dec. 3, 2013) as well as the ARMA's Generally Accepted Recordkeeping Principles: Principle of Disposition, <http://www.arma.org/r2/generally-accepted-br-recordkeeping-principles> (last visited Dec. 3, 2013).

30 For further information on legal holds, see *The Sedona Conference Commentary on Legal Holds: The Trigger & The Process*, 11 SEDONA CONF. J. 265 (2010), <https://thesedonaconference.org/download-pub/470>.

Disposition

Once the organization verifies that no legal, regulatory, or operational requirements apply to the information, disposition decisions can be made. In some circumstances, an organization may be able to determine from readily available information whether a record retention or legal preservation requirement applies. In other circumstances, a more detailed investigation and analysis may be required. The analytical approach to such situations is beyond the scope of this Commentary and is discussed more fully in the Sedona publication entitled, “*The Sedona Conference Commentary on Inactive Information Sources*.”³¹

Principle 7. When information governance decisions require an organization to reconcile conflicting laws or obligations, the organization should act in good faith and give due respect to considerations such as privacy, data protection, security, records and information management, risk management, and sound business practices.

Organizations often confront conflicting laws or obligations that apply to the same information, particularly when the organization conducts business across numerous jurisdictions.³² A common example involves the tension between the European Union Data Protection Directive, which prohibits transferring “personal information,” and United States federal court jurisprudence that mandates the production of such information during the discovery process.³³ In other circumstances, an organization may be required to preserve certain information for a specified period of time, while another jurisdiction may require such information be destroyed upon the owner’s request.

When faced with information governance decisions triggered by such conflicts, the organization’s key objective should be good faith compliance with all laws and obligations. Due deference should be afforded to conflicting laws or obligations, particularly when the conflict arises out of interests that span different jurisdictions.³⁴ Further, the most significant legal/regulatory and business considerations should be prioritized; not all conflicts are capable of complete resolution, and the organization will ultimately need to balance the competing needs, demands, and viewpoints of the stakeholders involved. To the extent compliance with all laws and obligations is not possible or practical; the organization should thoroughly document its efforts to reconcile the conflict and its resulting decision-making process.

31 See, *The Sedona Conference Commentary on Inactive Information Sources*, (2009) <https://thesedonaconference.org/download-pub/64>.

32 *Devon Robotics v. DeViedma*, Civil Action No. 09-cv-3552 2010 WL 3985877 (E.D. Pa. Oct. 8, 2010). The plaintiff in a breach of fiduciary duty and tortious interference requested all ESI relating to the former employee defendant, his Italian employer (a rival), and the alleged breach of contract between the plaintiff and the defendant’s new employer. The defendant moved for a protective order regarding the production of “documents owned by his employer,” arguing that the disclosure was prohibited by the Italian Personal Data Protection Code. The court found that the defendant did not show good cause for a protective order and denied the motion, writing that the defendant “made nothing but a blanket assertion that any disclosure could violate Italian law.” The court also stressed the importance of the requested ESI to the plaintiff’s claims and the comity factors outlined in *Societe Nationale* (482 U.S. 522 (1987)) weighed in favor of disclosure.

33 See, e.g. *Heraeus Kulzer, GmbH v. Biomet, Inc.*, 633 F.3d 591 (7th Cir. 2011).

34 For example, with respect to the transfer of information from France to the U.S. for use in legal proceedings, which allegedly would have violated a French blocking statute, the U.S. Supreme Court held that U.S. courts should “take care to demonstrate due respect for any special problem confronted by the foreign litigant on account of its nationality or the location of its operations, and for any sovereign interest expressed by a foreign state.” *Société Nationale Industrielle Aérospatiale v. United States District Court for the Southern District of Iowa*, 482 U.S. 522, 546 (1987). In so doing, “the concept of international comity requires in this context a ... particularized analysis of the respective interests of the foreign nation and the requesting nation.” *Id.* at 543-44.

Principle 8. If an organization has acted in good faith in its attempt to reconcile conflicting laws and obligations, a court or other authority reviewing the organization's actions should do so under a standard of reasonableness according to the circumstances at the time such actions were taken.

An organization's actions may be subject to review by a court or other governing authority regarding its attempt at resolving conflicting laws and obligations. That review should consider the specific circumstances when the information governance decision under review was made. Any judgment of the correctness of past actions to resolve conflicts should be based solely upon what was known at the time the decisions were made. Where a party has acted in good faith, it would be patently unfair to consider what they might have known had they possessed superior prescience.³⁵

Application of the reasonableness standards requires that a court or other authority objectively assess the organization's actions or decisions in comparison to the actions or decisions made by a hypothetical, similarly-situated organization acting reasonably under the same circumstances. In *Lewy v. Remington Arms Co., Inc.*, 836 F.2d 1104 (8th Cir. 1988), the court outlined factors to be considered in assessing the reasonableness of a record retention policy for a spoliation instruction, including: (i) whether the policy was reasonable considering the facts and circumstances surrounding the relevant documents (i.e., whether a three year retention policy is reasonable for a class of materials, such as email); (ii) whether any lawsuits relating to the documents had been filed, or may have been expected; and (iii) whether the document retention policy was instituted in bad faith. *Id.* at 1112.

In determining good faith, courts or other authorities should give due deference to decisions by corporate officers or directors by applying the "business judgment rule," which is a presumption that a business decision was made "on an informed basis, in good faith and in the honest belief that the action taken was in the best interests of the company." *Aronson v. Lewis*, 473 A.2d 805, 812 (Del. 1984) (citations omitted).

Principle 9. An organization should consider reasonable measures to maintain the integrity and availability of long-term information assets throughout their intended useful life.

If the intended useful life of an information asset is long enough that risks or concerns may arise regarding the ongoing integrity and availability of the information, then organizations should consider appropriate measures designed to protect those information assets. Therefore, *long-term* planning for availability and integrity depends on the circumstances involved, including the asset's purpose and storage media options.

For example, if your intended retention period is 25 years and the media format you will be using has an expected life of 12 years, then specific planning will be required to

35 *The Sedona Conference International Principles on Discovery, Disclosure & Data Protection*; Best Practices, Recommendations & Principles for Addressing the Preservation & Discovery of Protected Data in U.S. Litigation (European Union Edition), (2011), <https://thesedonaconference.org/download-pub/495>. Principle 2: "Where full compliance with both Data Protection Laws and preservation, disclosure, and discovery obligations presents a conflict, a party's conduct should be judged by a court or data protection authority under a standard of good faith and reasonableness." See also, ABA Resolution 103 (2012) (adopted), http://www.americanbar.org/content/dam/aba/administrative/house_of_delegates/resolutions/2012_hod_midyear_meeting_103.doc. 26k-2012-11-10: "[T]he American Bar Association urges that, where possible in the context of the proceedings before them, U.S. federal, state, territorial, tribal and local courts consider and respect, as appropriate, the data protection and privacy laws of any applicable foreign sovereign, and the interests of any person who is subject to or benefits from such laws, with regard to data sought in discovery in civil litigation."

ensure the ongoing integrity and availability of that information. Failing to ensure the integrity and availability of information assets may bring the risk of sanctions if an organization is unable to fulfill e-discovery obligations.³⁶

This principle is limited to “systems of record”, meaning that copies (such as convenience copies) are outside its scope. Backup and recovery, disaster recovery, and redundant storage paradigms such as ‘RAID’ are well-understood disciplines dictated by operational business continuity requirements and are therefore not covered by this Commentary. Logical defects prior to “long-term” storage also are not covered by this principle or Commentary.

Long Term Digital Assets

The phrase “long-term” is used to mean a time-frame sufficiently long to involve planning for concerns such as the physical degradation of the storage medium or the impact of changing technologies.

Planning for the ongoing integrity and availability of long-term information assets is important for both physical and digital information, but it is important for digital assets that may have a long lifecycle or retention period. The risks and considerations should be evaluated as part of the long-term retention strategy.

To maximize the probability of ensuring the ongoing integrity and availability of digital assets throughout their intended useful life, organizations should make a good-faith attempt to balance risk and cost. Creating a long-term retention strategy appropriate to the value and type of the information involves considering a broad range of factors pertaining to the digital assets and the circumstances of the organization itself. These factors should include business value, regulatory importance, intended retention schedule, legal hold status, file format, continued availability of the technologies required to access and read, the likely failure rate of the storage medium as it is configured, the available budget and resources of the organization, and/or (for 3rd party services such as *cloud* storage, SaaS, etc.), the contractual agreements between the customer and provider.³⁷

Principle 10. An organization should consider leveraging the power of new technologies in its Information Governance program.

For many organizations, reliance on end-users to effectively manage information continues to work well. These organizations should consider how technology can help individuals to better manage the information that they are responsible for, and to monitor management of the information. Examples of the former include limitations on the size of email accounts, or systems that automatically delete emails unless they are moved from the inbox or sent box. Appropriate use of this technology can significantly decrease the cost and risk of e-discovery because emails frequently make up a significant percentage of information that is collected for litigation or government investigations. Similarly, organizations should consider using technology that automatically deletes voicemails after a fixed number of days. Companies can also monitor for over-retention by providing management with lists of the largest email accounts or reports on data that has not been accessed recently.

³⁶ *United States v. Universal Health Servs., Inc.*, No. 1:07cv000054, 2011 WL 3426046 (W.D. Va. Aug. 5, 2011).

³⁷ For a more detailed explanation of the specific areas of risk for digital assets, see Appendix C.

However, organizations should consider using advanced tools and technologies to perform various types of categorization and classification activities. While the rapid advances in technology threaten to render obsolete the technology described in this commentary, an organization should consider using technologies such as machine learning, auto-categorization, and predictive analytics to perform multiple purposes, including: (i) optimizing the governance of information for traditional RIM; (ii) providing more efficient and more efficacious means of accessing information for e-discovery, compliance, and open records laws; and (iii) advancing sophisticated business intelligence across the enterprise.

Machine Learning, Auto-Categorization, and Predictive Analytics Defined

Machine learning is the “[f]ield of study that gives computers the ability to learn without being explicitly programmed.”³⁸ Training filters to recognize spam email is one common example of machine learning. In theory, just about any classification problem arising in information governance can benefit from being modeled by machine learning techniques. Some of these techniques do not rely on human intervention: for example, clustering or auto categorizing data into data types or classifications can be accomplished through software alone analyzing the properties of a data set.

One machine learning technique of particular utility involves active learning by software through human interaction on the front end, where humans train the systems to learn through examples. “Predictive coding” and “technology-assisted review” are terms used in the e-discovery arena that rely on humans coding seed sets of data into responsive and nonresponsive categories, with software then analyzing the remaining huge repositories of data.³⁹ As used here, “predictive analytics” means any machine learning technique that combines human intervention on the front end with the power of machine learning, to optimize the classification of information through automated rules.

New Technologies Meet Traditional RIM

If the structure or volume of information flowing through networks does not allow continued reliance on “end-users” to categorize content, organizations should consider taking steps that shift the burden of traditional records and information management from individuals to technology through auto-categorization of content. Organizations should, therefore, consider taking steps that shift the burden of traditional records and information management from individuals to technology through auto-categorization of content. For example, organizations may use existing software to analyze and categorize the contents of email for purposes of defensible deletion of transitory, non-substantive or non-record content.⁴⁰ Organizations increasingly utilize predictive analytics to assist in categorization functions, where individuals train software to differentiate between types of records.

For e-discovery, the first judicial opinions approving the use of predictive coding and technology-assisted review techniques for document review in e-discovery were published in 2012.⁴¹ In one case, the court stated that “the Bar should take away from this

38 Arthur L. Samuel, “*Studies in Machine Learning Using the Game of Checkers*,” IBM JOURNAL OF RESEARCH AND DEVELOPMENT 3(3):211-229 (1959).

39 See generally, The Grossman-Cormack Glossary of Technology Assisted Review, 7 FED. CTS. L. REV. 1 (2013).

40 The National Archives and Records Administration has endorsed the use of email archiving and capture technologies using smart filters to sort content through role-based and rule-based architectures. See NARA Bulletin 2013-02, *Guidance on a New Approach to Managing Email Records*, (Aug. 29, 2013), <http://www.archives.gov/records-mgmt/bulletins/2013/2013-02.html>.

41 See, e.g., *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012), approved and adopted in *Da Silva Moore v. Publicis Groupe*, No. 11 Civ. 1279(ALC)(AJP), 2012 WL 1446534 (S.D.N.Y. Apr. 26, 2012); *Global Aerospace Inc., et al. v. Landow Aviation, L.P., et al.*, No. CL 61040, 2012 WL 1431215 (Va. Cir. Ct. Apr. 23, 2012); *In re Actos (Pioglitazone) Products*, No. 6-11-md-2299, 2012 WL 3899669 (W.D. La. July 27, 2012).

Opinion ... that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review.”⁴² An important study by the Rand Corporation, anticipating this new direction in the law, concluded that predictive coding may significantly reduce e-discovery costs by reducing the number of documents requiring eyes-on review.⁴³

Predictive Analytics and Compliance

Predictive analytics is also increasingly being utilized by organizations outside of the e-discovery context, including in investigations and as an element of compliance programs. Predictive analytics is being used in compliance programs to predict and prevent wrongful or negligent conduct that might result in data breach or loss. Similar to how this technology is being used in litigation and investigations, predictive analytics is being used as an early warning system. To this end, companies use exemplar documents, sometimes in conjunction with search terms, to periodically search a target corpus of documents, usually email, to detect improper conduct.

Predictive Analytics and Business Intelligence

At its most fundamental level, predictive analytics assists in identifying information that may help to answer a question. There is no limit to the questions predictive analytics can help answer. Companies are beginning to use predictive analytics to develop business intelligence about the company, its information assets, and the market in which it operates.

Principle 11. An organization should periodically review and update its Information Governance program to ensure that it continues to meet the organization’s needs as they evolve.

Organizations and their environments change. The footprint and nature of the organization’s operations may expand, contract, or transform, and its technology capabilities and uses will evolve. The organization’s environment will also change, including legal requirements for the retention, protection, preservation, and disposal of information. And new information-related risks will also arise as time passes. Review of at least some aspects of many organizations’ Information Governance programs is legally required,⁴⁴ and regardless, is prudent given the inevitability of organizational and environmental change. Organizations, therefore, should periodically review and update their Information Governance program.

42 *Da Silva Moore*, 287 F.R.D. at 193.

43 N. Pace & L. Zakaras, “*Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*,” RAND Report (2012), <http://www.rand.org/pubs/monographs/MG1208.html>.

44 For example, HIPAA policies and procedures must be reviewed periodically and updated as needed in response to environmental or operational changes affecting the security of electronic protected health information. 45 C.F.R. § 164.316(b)(2)(iii). HIPAA security measures must also be reviewed and modified as needed to continue providing reasonable and appropriate protection for ePHI. 45 C.F.R. § 164.306(e). Comprehensive information security programs for customer information under the Gramm-Leach-Bliley Act must be evaluated and adjusted in light of any material changes in operations or business arrangements. 16 C.F.R. § 314.4(e). Entities subject to the FTC’s Red Flags Rule must ensure that their mandated Identity Theft Program is updated periodically to reflect changes in risks to customers or to their safety and soundness regarding identity theft. 16 C.F.R. § 681.1(d)(2)(iii). And entities that own or license personal information about Massachusetts’ residents must review their information security measures at least annually or whenever a material change in business practices reasonably implicates the security or integrity of records containing such personal information. 201 CMR. 17.03(2)(i).

Program review differs from the monitoring activities that should be embedded in the organization's Information Governance program. Such monitoring activities observe whether information-related practices comply with the program's rules and risk controls. *See* Principle 5, Accountability. The program review should seek to determine whether the program itself, and its rules and risk controls, remain appropriate for governing the organization's information in light of organizational and environmental changes. A flawlessly-executed Information Governance program will still result in compliance and risk exposures if elements of the program have become obsolete due to changed circumstances.

The review of the Information Governance program is akin to the assessment described under Principle 4. The organization should:

- identify any significant changes in its life cycle practices for information;
- identify significant changes in applicable compliance requirements and risks regarding its information;
- review the organization's strategic objectives for information governance in light of internal or external changes; and
- review the results from monitoring and measuring performance of the organization's Information Governance program, as an indicator of whether the program's rules and risk controls are adequate or should be refined.

Those responsible for administering the organization's Information Governance program should be involved in the program review. The need for objectivity in conducting such a review may make it valuable to have an independent review of the program. And ultimately, because senior leadership is responsible for the results of information governance at the organization, such senior leadership should participate appropriately in the review process, receive the results of the review, and then provide direction, support, and resources for needed changes in the program.

No bright-line rule governs how frequently an Information Governance program should be reviewed. As with other business-driven initiatives, the frequency of review will most likely depend on many factors relating to the organization.⁴⁵ If an organization is rapidly changing through frequent acquisitions and divestitures, or periodically undergoes major updates to its technology systems, then its information environment is likely to be ever-changing to adapt to its new structure or systems. Alternatively, if an organization is relatively mature, has a stable operations model, or is not governed by frequently changing governmental regulations, it may be reasonable for it to conduct its reviews less frequently (i.e., biannually), to reassess and identify potential modifications to its recordkeeping, data security, and operational requirements. Further, an organization may be subject to external pressures, such as regulations subject to frequent modification or regular compliance audits that require systemic changes; in such cases, the organization should be prepared to review and revise its information governance policies on an ongoing basis to meet the challenges posed by such changes. An organization should track pending legislation and regulations

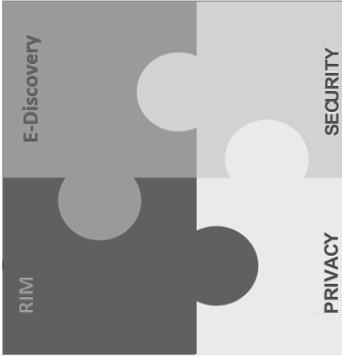
⁴⁵ Determining the appropriate frequency of review is a matter of business judgment. Courts generally defer to decisions by corporate officers and directors pursuant to the "business judgment rule," which is built upon the presumption that business decisions are made "on an informed basis, in good faith and in the honest belief that the action taken was in the best interests of the company." *Aronson v. Lewis*, 473 A.2d 805, 812 (Del. 1984), (*overruled on other grounds by Brehm v. Eisner*, 746 A.2d 244 (Del. 2000)).

relevant to its industry to facilitate continued compliance with the regulations that affect its operations. It would be prudent to include a review of its information governance policies and procedures as part of its response to such developments.

Because of the ongoing program review, update, and execution, an organization will have reasonable assurance its Information Governance program continues to meet both legal requirements and also the organization's strategic objectives for information.

APPENDIX A

Intersections



Intersections Create Opportunities and Challenges

Although the functional areas of RIM, E-Discovery, Privacy and Information Security are frequently separate, a successful Information Governance program requires them to work together. As there is some natural overlap between the three groups, some of this will come naturally and provides opportunities to combine resources and budgets. Conversely, in some areas the goals of intersecting groups may clash and require resolution before an initiative can move forward. Identifying and leveraging these areas early in a program is an important task. The table below defines many of the synergies and conflicts in the intersections of these groups.

<p>Functional Area Focus</p> <p>RIM <i>Primary Focus:</i> <i>RIM programs ensure that records and information are properly maintained, accessed, and ultimately disposed of in accordance with statutory and regulatory requirements and with consumer expectations. They also ensure that those organizations with which there is a third-party relationship endorse the same safeguards and have appropriate means of guaranteeing compliance.</i></p>	<p>RIM Intersection with Functional Area</p> <p>N/A</p>	<p>E-Discovery Intersection with Functional Area</p> <p>Potential Synergy:</p> <ul style="list-style-type: none"> • Similar metadata concerns. • Work together to respond to document requests by locating and preserving relevant information. • Support consistent defensible disposition of information in accordance with an organization's legal, regulatory and operational requirements. • Enables organization to know what they have and identify, preserve, retrieve, search, produce and appropriately destroy in normal course of business. • RIM protects against loss of content that could lead to sanctions, financial loss and brand risk during e-discovery. • RIM serves as evidence of official policy and helps ensure that evidence can be authenticated. <p>Potential Friction:</p> <ul style="list-style-type: none"> • Could be responsible for retention of drafts or outdated content due to relevancy. • RIM focus is can be more narrowly targeted to "records" while e-discovery is broadly on ESI 	<p>Privacy Intersection with Functional Area</p> <p>Potential Synergy:</p> <ul style="list-style-type: none"> • Defines requirements for identification and classification of sensitive information. <p>Potential Friction:</p> <ul style="list-style-type: none"> • RIM may need wide access and distribution while Privacy seeks limits. 	<p>Security Intersection with Functional Area</p> <p>Potential Synergy:</p> <ul style="list-style-type: none"> • Ensures that sensitive information is properly maintained, identified and content is classified. • Ensures that sensitive data and information is properly maintained, accessed and disposed of according to legal and regulatory requirements. <p>Potential Friction:</p> <ul style="list-style-type: none"> • RIM may need wide access and distribution while Security seeks limits. • Encryption may be required in Security but frustrate accessibility by RIM.
--	--	---	---	---

Functional Area Focus	RIM Intersection with Functional Area	E-Discovery Intersection with Functional Area	Privacy Intersection with Functional Area	Security Intersection with Functional Area
<p>E-Discovery</p> <p><i>Primary Focus:</i> <i>Preservation of electronically stored information that is potentially relevant to impending or ongoing litigation and is processed in a timely, auditable and efficient manner.</i></p>	<p>See RIM / E-Discovery intersection above</p>	<p>N/A</p>	<p>Potential Synergy:</p> <ul style="list-style-type: none"> • Identification at point of creation of information subject to privacy regulations may reduce risk that private information will be produced. <p>Potential Friction:</p> <ul style="list-style-type: none"> • Producing private information protected by another country's laws can result in criminal or civil sanctions. • Refusing to preserve and produce private information may result in civil or criminal penalties under US Law. 	<p>Potential Synergy:</p> <ul style="list-style-type: none"> • Ensures that sensitive data and information is available, if relevant; and that out-of-date information is disposed of according to legal and regulatory requirements. • Satisfies an organization's 'duty to preserve' for forensic collections. <p>Potential Friction:</p> <ul style="list-style-type: none"> • Security encryption requirements can hamper e-discovery efforts.

Functional Area Focus	RIM Intersection with Functional Area	E-Discovery Intersection with Functional Area	Privacy Intersection with Functional Area	Security Intersection with Functional Area
<p>Security</p> <p><i>Primary Focus:</i> <i>Ensuring the confidentiality, integrity, and availability of information and assets.</i></p>	<p>See RIM/Security intersection above</p>	<p>Potential Synergy:</p> <ul style="list-style-type: none"> Ensures that sensitive data and information is available, if relevant, and that out-of-date information is disposed of according to legal and regulatory requirements. Satisfies an organization's "duty to preserve" for forensic collections. <p>Potential Friction:</p> <ul style="list-style-type: none"> Security encryption requirements can hamper e-discovery efforts. 	<p>Potential Synergy:</p> <ul style="list-style-type: none"> Security enforces the access rights defined by privacy. <p>Potential Friction:</p> <ul style="list-style-type: none"> Privacy requirements may hamper security investigations 	<p>N/A</p>

APPENDIX B

Maturity Continuum as it Relates to Independence

It is important to consider the independence of the Information Governance function of an organization when making determinations such as assessing the current maturity, or planning how to increase the future maturity of an Information Governance program.

While not all organizations have a sufficiently mature Information Governance program to warrant the appointment of a C level executive in this role, we believe that organizations must ultimately view information governance as requiring an executive leader that is accountable to the CEO or COO in order to ensure that decisions are made in the best interests of the overall organization, rather than for the good of discrete departments.

A common difficulty when balancing costs and risks occurs when the choices have dissimilar characteristics that make comparison difficult. For example, a clearly-defined cost saving may need to be weighed against a high impact, low-probability event, such as statutory fines in the event of leakage of protected data, where it is difficult to quantify the probability of the event occurring or the costs. Whatever risk management methodology is used to balance cost and risk, it will be more accurate to make the determination by looking at the problem from the perspective of the overall organizational impact.

However, if the executive in charge of information governance reports to an individual department, there is the potential for the interests of that department to be given greater weight than the overall interests of the organization. The simple fact that the department to which the executive reports funds their work and rates their job performance may result in such a bias.

Therefore, the level of independence of the information governance function of an organization is an important component of the information governance maturity continuum.

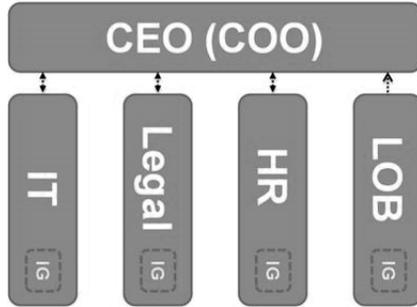
Maturity and Independence

The following discussion is intended as a reference to aid in assessing the current level of maturity of an information function, planning how to move an organization further along the information governance maturity continuum, or making a determination as to what is *sufficient* independence for a given organization. The concepts described below can be adapted for the specific circumstances of an organization.

Note: The following graphics are highly simplified, generic representations of potential organizational structures at varying points along the maturity continuum. The graphics depict the coordination and accountability at a departmental level. Specific functions such as RIM, Privacy, Information Security, E-Discovery, etc. are intentionally not shown because they generally reside within a stakeholder department.

Immature

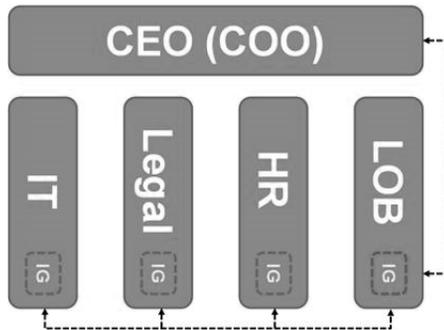
Immaturity is characterized by a lack of over-arching coordination of information governance stakeholders and no single point of accountability to the CEO or COO for overall governance of information.



At the immature end of the maturity continuum, lack of coordination creates a potential for important requirements being missed. Decisions and requirements reside in silos, and cross-functional coordination is ad hoc. There is a potential for departmental decisions that conflict with other stakeholder requirements and which are not in the interests of the organization overall. There is also a potential for inconsistent treatment of different items in the same category in the same circumstances.

Less Mature

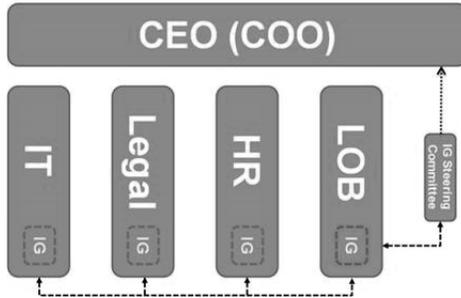
At this area of the maturity continuum, ownership of information governance process resides within a stakeholder department.



There is a potential conflict of interest since ownership must reside in a stakeholder department, which presents the problem of misaligned incentives.

More Mature

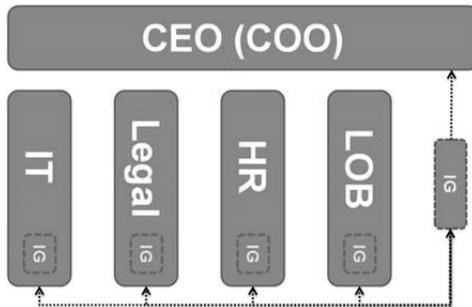
At this area of the maturity continuum, ownership of Information Governance process resides in a stakeholder department but is accountable to a steering committee of C level executives from the stakeholder departments who are accountable to the CEO or COO.



There is still a potential for conflict of interest for the executive in charge of Information Governance (who resides in a stakeholder department) and for the C level executives on the Information Governance steering committee because the goals of the individual departments may conflict with the goals of the overall Information Governance program.

Mature

A mature Independence Governance function is characterized by an executive who resides in a separate Information Governance department who is accountable to the CEO or COO for coordinating stakeholders across all departments and functions and balancing decisions for the benefit of the organization overall.



APPENDIX C

Risks Associated with Digital Assets**Risks**

There are specific areas of risk for digital assets that organizations should consider, including:

Integrity

The term “integrity” is used to mean the authenticity and reliability of the information. In some situations this may simply mean the logical content of the information has not been altered. In other situations it may mean the file can be guaranteed not to have changed.

The integrity of the information, or of information required to access the information (such as an index or necessary metadata) may be compromised by factors such as unauthorized alteration, or degradation of the storage medium. These risks can become particularly acute during platform migration.

Consideration should be given to: (a) the level of integrity required both for the digital asset in question and the technologies required to read and access the data, and (b) the level of difficulty involved in repairing or recovering damaged digital information.

Careful consideration should be given to the file format, storage medium (including the configuration of that storage medium), and the circumstances of operation and storage, in order to ascertain the likelihood of data loss.

Digital storage media without moving parts such as flash drives, solid state drives, and tape, or with rarely moving parts (such as storage devices intended for infrequent use that power off when not in use) still fail. Unused storage media on a shelf (for example, forensic collections on individual storage media in an evidence lab) will eventually become unusable. Given the relatively short lifespan (say, three-to-five years) of some items of storage media, a legal hold or retention requirement that may potentially exceed the reasonably expected lifespan could necessitate specific *long-term* planning due to the failure rate of the technology involved.

Availability

The term “availability” is used to mean “able to be used when needed,” which includes:

- any element (such as security mechanisms to protect the data, access rights required to access the data, or applications required to interpret or read the data);
- being able to access information in a timely manner (for example within applicable service-level agreements, contractual requirements, or timeframes indicated by legal requirements);

- being available within a pre-agreed lead-time (depending on business need – for example, a week).

Note that availability does not necessarily mean continuous availability.

The availability of information, or information required to access the information (such as an index or necessary metadata) may be compromised by obsolescence or unavailability of technology required for accessing the information (or index, or necessary metadata) in a timely manner.

Considerations

When planning for ongoing integrity and availability of digital assets throughout their intended useful life, important considerations include:

Technology Refresh Period

The phrase “technology refresh period” is used to refer to the timeframe in which technology components are expected to fail, and within which planning needs to occur for replacing those components.

Organizations should exercise prudence when considering the technology refresh period for long-term digital assets. For example, if the expected lifespan of the storage medium is seven years, then the technology refresh period should be less than seven years. The timing of the technology refresh period compared to the technology’s expected lifespan is a matter of risk calibration and business judgment.

Planned Migrations

Obsolescence of technology is a major consideration in long-term storage of digital assets and requires careful planning. Migrations (moving to a new platform for the archive as a whole or for a component of the archive) are a consequence of obsolescence that must be planned. All elements of the archiving system including search-and-retrieval capability as well as storage medium should be considered in terms of obsolescence. Organizations should consider creating an obsolescence review period as part of their long-term archival planning, because unlike a technology refresh period (which can be ascertained in advance for each technology refresh cycle by reference to the expected life of the technology components) the probable time of obsolescence may not be knowable in advance.

Migrations may also require format conversions, and integrity-checking technologies (see below) are particularly critical to ensure the data is not inadvertently changed during a migration.

Matching Storage Medium to the type of Electronic Information

It is important to match the characteristics of the storage medium to the requirements of the information being stored. For example, micrographics work particularly well for text documents – particularly text documents held for reference purposes – but not for binary files such as audio files or CAD (Computer Aided Design) files. Micrographics also may not work well for files that need to be in digital format when used because a scanning or conversion process will be required before the file can be used.

The expected failure rate of the storage medium should be considered in terms of the expected retention period. Regulated utilities or pipelines often involve document retention periods of decades, sometimes over 50 years, often longer than the life of the plant.

Integrity-Checking Technologies

Passive integrity-checking technologies can be used to assess if a file has changed. These technologies include such mechanisms as hash values created by hash algorithms computed when a file is retrieved and if the file has changed. Unfortunately, passive integrity-checking technologies have no inherent mechanism to repair files and restore them to their original form; they can only alert you to the fact that a problem has occurred.

Active integrity-checking technologies can be used not only to assess if a file has changed but also (if appropriately configured) to restore a file to the original form as when it was stored. There are many proprietary examples of integrity-checking archive technologies. Because these technologies are generally well-understood and well documented, they are not discussed further here.

Long Term Physical Information Assets

When considering storage using physical mediums such as paper, it is important to ensure that the expected life of the storage medium exceeds the retention requirements. In the case of printed paper, the expected life of different types of paper, as well as different types of ink, can vary a great deal. It is also important to consider the storage conditions (such as humidity and temperature) required to ensure the ongoing integrity of the physical assets because this can affect the expected life of the physical storage medium.

APPENDIX D

The Quantitative/ROI Business Case

As discussed in the Commentary, a successful information governance approach requires both strategic commitment (adoption as an organizational priority) and tactical efforts. This Appendix discusses approaches to establishing an acceptable ROI for particular projects.

A typical ROI analysis weighs the benefits of a particular project against its cost, and calculates the length of time it will take to recoup the cost. The quantitative aspects of the business case are best determined by focusing on specific applications of information governance to identified problems or opportunities, or to discrete projects for implementation of the Information Governance program.¹

The quantifiable benefits from pursuing information governance generally fall into four main categories: optimizing corporate value, risk reduction, hard cost avoidance, and soft cost avoidance.

Optimizing Corporate Value

Information governance can help make information assets available for new, valuable uses. It can also allow organizations to derive value from engaging in what might otherwise be cost-prohibitive endeavors, due to efficiencies and cost savings realized through information governance practices. In general, Gartner has identified the following as possible “adds” to corporate value from an Information Governance program:

- **Effectiveness:** Such as due to document-centric collaboration tools;
- **Cost/efficiency:** For example, from imaging/workflow solutions to replace traditional paper-oriented processes;
- **Customer service:** Such as from customer-relationship solutions that lead to better market penetration and customer satisfaction;
- **Competitive advantage:** As more modern tools and reliable information allows for speedier delivery of goods or services to customers; and
- **Revenue:** Such as a result of enhanced social media and web presences and solutions.²

By way of example, a core benefit of an Information Governance program is to ensure that information used for different purposes across the enterprise – e.g., for sales and marketing, but also for planning, billing, fulfillment, financial, customer feedback and other downstream purposes – is reliable or trustworthy, accurate,

¹ See generally, S. Soares, *Selling Information Governance to the Business: Best Practices by Industry and Job Function* (2011) (providing insight into the best ways to encourage businesses to implement an information governance program).

² See Gartner, “*First 100 Days: Enterprise Content Management Initiatives*” (July 7, 2011), available at <http://www.gartner.com/id=1739415>.

and in formats usable across platforms or applications. Achieving these objectives requires that IT understand not only the business purposes and objectives but also whether data elements require special protections or treatments (e.g., for legal, RIM, privacy or security reasons).³ Yet, oftentimes when a large organization initiates such a program, it finds that different business units or functions use different terminology for the same content concept. For example, an organization may refer to outside business partners as *vendors*, *suppliers*, *associates*, or *providers* and collect various information about such entities in systems that support particular functions within the organization. But if the terminology – or application – differs between and among business units, opportunities to cross-sell or otherwise leverage the information about the business partners may be missed.⁴ Thus, an early goal for an Information Governance program may be to develop a common vocabulary and understanding of what information-related assets exist; once that is done, the organization may realize that business advantages may be achieved – at virtually no cost – by cross-utilizing existing information or systems.⁵

Mergers and acquisitions, or technology upgrades, also present opportunities (and challenges) for improving data quality and corporate revenues by, for example, merging (and purging) customer lists to identify strong customers across multiple business lines.⁶

Risk Reduction

Risk reduction is also a significant benefit of information governance. Business value may not be realized if an unanticipated risk creates an unexpected cost. For example, organizations may leverage information over the short-term (e.g., email for current communications), but once the information is no longer useful, the ESI is often stored away, rarely accessed, and often never re-assessed to determine whether the benefits of continued retention outweigh the risks. Thus, what was once a business asset may become a source of risk for certain organizational areas such as compliance or e-discovery, while providing little or no benefit for other organizational areas such as business units. Through proper information governance, organizations can recognize these perils and elect to remediate the un- or under-utilized information assets, and optimize the business value of information while managing the associated risks.

Many types of adverse events can be avoided through effective information governance. The value of risk reduction can be estimated by quantifying the potential losses that would result if an adverse event occurred and determining the reduced likelihood of such an occurrence due to effective information governance. Some examples of risks posed by information assets follow:

³ See, e.g., Soares, *supra*, at 149.

⁴ As another example, it has been reported that one manufacturing company discovered and eliminated 37 unique definitions of “customer” across its enterprise, and agreed on a single, standard definition. Robert Routzahn, “*Business and IT Collaboration: Essential for Big Data Information Governance*,” IBM Data Magazine, (July 5, 2013), <http://ibmdatamag.com/2013/07/business-and-it-collaboration-essential-for-big-data-information-governance/>.

⁵ See, e.g., “*The Sedona Conference Commentary on Finding the Hidden ROI In Information Assets*,” The Sedona Conference, (Feb. 2011), <https://thesedonaconference.org/download-pub/466>.

⁶ A medical device manufacturer estimated that improving ship-to addresses in a 100,000 item database could increase aftermarket sales by \$1 million. Soares, *supra*, at 69.

- a. **Data Leakage:** Many companies have valuable intellectual property that is more likely to be lost or leaked to the public and/or competitors if not properly managed through policies and procedures that emanate from a mature Information Governance program.
- b. **Privacy Breaches:** A myriad of regulations applicable to particular sectors in the U.S. (e.g., HIPAA to health information, GLBA to financial institutions, PERPA to federally funded educational institutions) require certain data to be protected and impose fines and other sanctions when the data is not properly protected or is improperly disclosed.
- c. **Security Lapses:** Regulations such as the self-regulatory Payment Card Industry Data Security Standards require companies to protect credit card and other payment information, or face fines.
- d. **Brand Impact:** A breach of private customer information, such as contact information or social security numbers, can adversely impact a company's brand and result in lost sales and/or consumer goodwill.
- e. **Litigation/Regulatory Risk:** Access to the most relevant information at the inception of litigation or a regulatory inquiry may allow for an earlier and more accurate assessment of litigation risk, and thus, permit such events to be more effectively and economically managed.

Hard Cost Avoidance

Many benefits flowing from an information governance initiative are based on the premise that certain future costs can be delayed, reduced or avoided entirely because lesser volumes of data will be kept in a more efficient manner. These benefits can be quantified, and in an information governance initiative, often arise from the following areas:

- a. **Storage:** Storage and maintenance costs can be radically reduced by the rationalizing data storage options, eliminating outdated ESI that no longer serves a legitimate business, legal or regulatory purpose, and moving valuable information that is occasionally and non-critically accessed to cheaper storage. A systematic approach to information governance may allow an organization to archive its less active and less critical data on less expensive tiers of storage, which in turn can eliminate unnecessary duplication of documents, associated backup overhead and better enable data disposition in line with organizational policy.
- b. **Outdated Backup Media:** Eliminating the retention of large (and outdated) quantities of backup media, such as magnetic tapes, reduces the costs of backup media and related storage, labor and transfer expenses.
- c. **Personnel Costs:** A successful Information Governance program will reduce the volume of ESI and make it easier to manage and to find information. Accordingly, fewer personnel would be required to manage the reduced volume, allowing the organization to realign resources appropriately.

- d. **E-Discovery Costs:** A reduced volume of electronic information can, in the event of litigation, reduce litigation costs *significantly*, because there will be less information to process and review.⁷

Soft Cost Avoidance

Other benefits resulting from improved information governance save time and effort that can be deployed for other activities. For example, having a more efficient method for storing and accessing email messages might save 30 minutes per day for each employee, netting a direct financial savings to the organization, or allowing employees to focus on more useful activities. Soft costs are often difficult to quantify, but the following are useful considerations:

- a. **Economies of Scale:** Managing information on an *ad hoc* basis can result in requirements and risks being overlooked, benefits not being realized, and tremendous amounts of inefficiency due to the redundancy of effort this entails. Economies of scale can be realized by having an over-arching Information Governance program at an organizational level, which generates processes and procedures to govern how ESI is handled.
- b. **Organizational Inefficiencies:** Organizations with excessive amounts of uncategorized ESI are often unable to locate needed information in a timely and efficient manner. An Information Governance program that creates an infrastructure for information assets promotes shorter client response times, allows the re-purposing of institutional knowledge, and enhances continuous improvement efforts.

⁷ A recent Rand survey states that the review process alone averages \$18,000 a gigabyte, meaning that with collection, preservation, hosting, etc., e-discovery costs can easily exceed \$20,000 a gigabyte. Pace, Nicholas M. and Laura Zakaras. *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*. RAND Corporation, (2012), <http://www.rand.org/pubs/monographs/MG1208>. Also available in print form.

THE SEDONA CONFERENCE DATABASE PRINCIPLES ADDRESSING THE PRESERVATION AND PRODUCTION OF DATABASES AND DATABASE INFORMATION IN CIVIL LITIGATION*

*A Project of The Sedona Conference Working Group on
Electronic Document Retention & Production (WG1)*

Author:
The Sedona Conference

2011 Editor-in-Chief:
Conrad J. Jacoby

2014 Editor-in-Chief:
Sherry B. Harris

2011 Editors:
David J. Kessler
Catherine L. Muir

2014 Editors:
David J. Kessler
Catherine L. Muir
Christopher Paskach

Thanks go to all who participated in the dialogue that led to this Commentary. We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group I. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.*

PREFACE

Welcome to the 2014 Edition of *The Sedona Conference Database Principles Addressing the Preservation and Production of Databases and Database Information in Civil Litigation*, another major publication of The Sedona Conference Working Group Series (“WGS”). This document contains numerous changes from the 2011 public comment version. The changes reflect the informal and formal suggestions and comments we received in the past few years. In addition, the changes take into consideration the continued evolution of law and best practices in the area over the past few years. The principles and accompanying text have been revised to harmonize the enhanced understanding of the technical, process, and legal issues that have emerged since publication for public comment.

The Sedona Conference Working Group on Electronic Document Retention and Production (WG1) recognizes that disputes over the discovery of electronically stored information in searchable data repositories are increasingly common in civil litigation. We hope this publication will provide practice guidance and recommendations to both requesting and producing parties and will simplify discovery in civil actions involving databases and information derived from databases.

The Sedona Conference thanks the drafting team and all WG1 members whose comments contributed to this publication. Special acknowledgement goes to David J. Kessler, Catherine L. Muir and Chris H. Paskach who assumed leading roles in revising the public comment version and resulting in the 2014 Edition. WG1 Steering Committee Liaison Sherry B. Harris provided a fresh perspective and an independent review of the publication. WG1 member Tim Hart provided thoughtful, substantive comments and suggested revisions to the public comment version, many of which were extremely valuable during the editing process.

We hope our efforts will be of immediate and practical assistance to judges, parties in litigation and their lawyers and database management professionals. We continue to welcome comments for consideration in future updates. If you wish to submit feedback, please email us at info@sedonaconference.org.

Craig W. Weinlein
Executive Director
The Sedona Conference
September 2014

TABLE OF CONTENTS

Executive Overview.....	175
The Sedona Conference Database Principles.....	176
I. Introduction.....	177
A. How Do Databases Differ from Other ESI?.....	177
B. Components of a Typical Database System.....	179
C. Assessing Relevance for Databases and Database Records.....	180
D. Preservation of Databases.....	180
E. Collecting and Producing Database Information.....	181
F. Potential Use of Database Information by a Requesting Party.....	182
G. Locating Specific Database Information through Queries.....	184
H. Databases and Database Information in a Third Party’s Custody or Control.....	185
II. Application of the Existing <i>Sedona Principles</i> to Databases and Database Information.....	186
A. Sedona Principle 3: The Early “Meet and Confer”.....	186
1. Redactions, Omitted Data Fields, and the Inadvertent Production of Privileged and Other Protected Data.....	186
2. Use and Role of Consultants and Technology Partners.....	187
3. Impact of Remote Jurisdiction and Location.....	188
B. Sedona Principle 5: Duty of Preservation.....	189
1. Burden of Preservation.....	189
2. Inventory and Default Retention Periods.....	190
C. Sedona Principle 6: Responsibilities of Responding Parties.....	191
1. Parties Must Understand Important Database Characteristics.....	191
2. The Responding Party Ordinarily Should Determine the Best and Most Reasonable Way to Identify, Extract, and Produce Relevant Data from Databases.....	192
3. Parties Must Consider the Database as It Is, Not as It Could be.....	193
4. Direct Examination of Databases.....	194
5. Documentation and Validation of Database Collections.....	195
6. Features and Limitations of the Technology and Tools that can be Applied to Databases to Identify and Extract Relevant Information.....	195

- D. Sedona Principle 12: Form of Production and Metadata196
 - 1. Mismatch of “Native Format” to Most Database Productions196
 - 2. Use of Standard Reports to Produce Database Information197
 - 3. Use of Fielded Tables to Produce Database Information198
- III. The Sedona Conference Principles for the Preservation and Production of Databases and Database Information (“The Database Principles”).....199
 - 1. Scope of Discovery199
 - 2. Accessibility and Proportionality205
 - 3. Use of Test Queries and Pilot Projects209
 - 4. Validation210
 - 5. Data Authenticity and Admissibility212
 - 6. Form of Production214

EXECUTIVE OVERVIEW

The Sedona Conference Working Group on Electronic Document Retention and Production has developed Principles addressing the preservation and production of databases, *The Sedona Conference Database Principles*. In these *Database Principles*, we offer a number of practical suggestions that we believe clarify the obligations of both requesting and producing parties and simplify discovery in matters involving databases and information derived from databases. We recognize that the specific facts of a litigation matter, combined with the implementation of relevant databases likely will raise additional retention and production issues not explicitly covered by these Database Principles. Even so, we believe that the groundwork laid by the Database Principles will provide valuable guidance to litigants facing novel issues of database retention and production. It is important to set reasonable expectations for the production of database information, and thus, an overarching theme of these Principles is that communication – between database management professionals and the attorneys who are asking them to search and export litigation-specific information, as well as between requesting and producing attorneys – is critical when working with databases. Many common disputes about issues such as the production format of data can be reduced or even eliminated through better dialogue between litigants.¹ We also find that better communication naturally will reduce “blunderbuss” requests for databases that typically encompass irrelevant or inappropriate information, or the production of terabytes of useless, undifferentiated data.

Our Commentary is divided into three discrete sections. Following a brief Introduction in Section I to databases and database theory, Section II addresses how The Sedona Principles, which pertain to all forms of ESI, may be applied to discovery of databases. Section III proposes six Principles that pertain specifically to databases and provides commentary to support our recommendations.

As database technology continues to evolve, we acknowledge that The Sedona Conference Database Principles will need to be revisited regularly to ensure that their guidance remains topical. At the same time, we believe that the Database Principles lay a foundation that will be valid both today and in the future for developing effective and practical solutions in this sophisticated area of the law.

¹ See *The Sedona Conference Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009 Supp.), available at <https://thesedonaconference.org/download-pub/3802>.

THE SEDONA CONFERENCE DATABASE PRINCIPLES

The Sedona Conference Working Group on Electronic Document Retention and Production (WG1) has been studying issues about the discovery of database information in civil litigation and has developed the following Principles addressing the preservation and production of databases, The Sedona Conference *Database Principles*.

1. **Scope of Discovery**

Absent a specific showing of need, a requesting party is entitled only to database fields that contain relevant information, and give context to such information, and not to the entire database in which the information resides or the underlying database application or database engine.

2. **Accessibility and Proportionality**

Due to differences in the way that information is stored or programmed into a database, not all information in a database may be equally accessible, and parties should therefore apply proportionality to each component of a database to determine the marginal value of the information to the litigation and the marginal cost of collecting and producing it.

3. **Use of Test Queries and Pilots**

Parties should use objective information, such as that generated from test queries, pilot projects, and interviews with persons with relevant knowledge to ascertain the burden and benefits to collect and produce information stored in databases and to reach consensus on the scope of discovery.

4. **Validation**

A responding party should use reasonable measures to validate that its collection from the database is both reasonably complete and did not inadvertently modify the ESI.

5. **Data Authenticity and Admissibility**

The proper validation of collection from a database does not automatically make the substantive information stored in the database authentic, admissible or true. These are separate issues that need to be analyzed by the appropriate decision makers.

6. **Form of Production**

The way in which a requesting party intends to use database information is an important factor in determining an appropriate format of production.

I. INTRODUCTION

Disputes over the discovery of information stored in databases are increasingly common in civil litigation. Part of the reason is that more and more enterprise-level information is being stored in searchable data repositories, rather than in discrete electronic files. Another factor is that the diverse and complicated ways in which database information can be stored has made it difficult to develop universal “best-practice” approaches to requesting and producing information stored in databases. The procedures that work well for simple systems may not make sense when applied to larger server-based systems. Similarly, data retention policies vary widely for different types of databases, from very short life-spans of data that can be measured in minutes or seconds to indefinite retention. (It is not uncommon for databases to have no purge or delete routines).

A. How Do Databases Differ from Other ESI?

Successfully working in a discovery context with databases and the structured data found in them requires a basic understanding of this form of electronically stored information (“ESI”) as it functions in the ordinary course of business.

Databases² generally contain “structured data,”³ rather than “unstructured data.” Structured data tends to have the following characteristics:

- Logical entities⁴ are decomposed into their constituent data elements (known as *fields* or *records*) at a highly granular level;
- Individual data elements are stored in specific assigned logical and physical areas within a series of files (or a single fielded table or a text delimited file⁵);
- These data elements are linked to each other by internal mechanisms, interpretable by the database software;
- These links or relationships may involve metadata elements stored within the database, in addition to the data elements of the logical entity; and
- Once properly assembled and formatted (e.g., in the form of a report), structured data is often readily understandable.

For example, in the case of a simple invoice being stored in a relational database, the logical entity “invoice” might consist of customer name, customer address, item ordered, cost of item, etc. These data elements themselves consist of more granular data elements. For example, customer name could be further decomposed into customer first name, customer middle initial, and customer last name. Similarly, item ordered could be

2 *The Sedona Conference Glossary: E-Discovery & Digital Information Management*, (4th ed.), 15 SEDONA CONF. J. 305 (2014) (“*The Sedona Glossary*”), defines a database as: “A set of data elements consisting of at least one file, or of a group of integrated files, usually stored in one location and made available to several users. ... Computer databases typically contain aggregations of data records or files....”

3 *The Sedona Glossary* defines structured data as: “Data stored in a structured format, such as databases or data sets according to specific form and content rules as defined by each field of the database.”

4 *The Sedona Glossary* defines logical entity as: “An abstraction of a real-world object or concept that is both independent and unique. Conceptually, a logical entity is a noun, and its relationships to other entities are verbs. In a relational database, a logical entity is represented as a table. Attributes of the entity are in columns of the table and instances of the entity are in rows of the table. Examples of logical entities are employees of a company, products in a store’s catalog, and patients’ medical histories.”

5 *The Sedona Glossary* defines text delimited file as: “A common format for structured data exchange whereby a text file contains fielded data where the fields are separated by a specific ASCII character and also usually contain a header line that defines the fields contained in the file.”

decomposed into item description, SKU number, and price. These data elements are commonly placed in structures called “tables,” which are used to organize the information, as defined further below.

By contrast, “unstructured” data⁶ tends to have the following characteristics:

- “Stand-alone” ESI consisting of a self-contained file or document (examples include MS Word, MS Excel, Adobe PDF, etc.);
- Generally does not require any highly technical knowledge to understand or use an individual file or document containing unstructured data; and
- Both the creation or selection of information to be included in the file or document and the way that information is formatted for display are left to the discretion of the creator of the file or document containing unstructured data.

Structured data may be found in contexts that you might otherwise expect to contain unstructured data, such as email database systems⁷ or websites (e.g., Lotus Notes, or WordPress). Conversely, unstructured data from time to time embedded in structured data (e.g., a customer invoice might be stored in a database column as a .pdf file). Both of these situations are outside the direct focus of this Commentary.

For structured data in a database, individual data elements or fields – each of which needs be accessed separately for relevance – must be assembled and viewed in context to be understood. Databases, however, impose strict rules that define how information can be entered, stored, and retrieved. For example, a particular database might store a customer’s name, John Q. Smith, as three discrete elements – first name (John), middle initial (Q), and last name (Smith) – each in separate data fields. Unlike the unstructured file, these separate elements must reference each other to be recalled and displayed as a whole name. Each database may have its own unique rules for storing and recalling elements of information. Additionally, different applications (even those written on the same type of database system) may be designed differently and may store a whole name (for example John Q. Smith) in a single field without dividing it further.

End-users commonly think of database information in terms of records they query, retrieve, and view. Although a database record may be the closest intellectual analog to a “document” within a database, records consist of separate data elements that may be stored in a number of ways within a database, such as in multiple tables, or across multiple databases. Thus, a “record” may not exist until actions by a user instruct the database application to assemble specified fields for display. Accordingly, a database record is not always an appropriate granular level of information to respond to a discovery request. At various times, key information may be found in a single data field, in a record made up of a set of selected fields, in a table containing a pool of records, or in a report that extracts discrete fields of information from multiple tables. Thus the extraction of responsive information from databases may often require specialized business or technical knowledge.

⁶ *The Sedona Glossary* defines unstructured data as: “[F]ree form data which either does not have a data structure or have a data structure not easily readable by a computer without the use of a specific program designed to interpret the data; created without limitations on formatting or content by the program with which it is being created. Examples include word processing documents or slide presentations.”

⁷ Although the email message content itself is unstructured, emails are accompanied by metadata in assigned fields, including, but not limited to, the sender, recipient, date, and time. The message content and metadata elements are stored together in an email database system, comprising an email record. The email database system stores individual email records, imposing the same storage format across all individual email records.

For instance, using the simple example of the organized collection of customer invoices, the customer “record” might be defined as a set of “fields,” composed of the following fields:

FIRST NAME:
 MIDDLE INITIAL:
 LAST NAME:
 STREET ADDRESS:
 SUITE NUMBER:
 CITY:
 STATE:
 ZIP CODE:
 TELEPHONE NUMBER:
 FAX NUMBER:
 EMAIL ADDRESS:
 COMMENTS:

Hundreds or thousands of such customer records may be stored in the database, with the elements for each customer arranged in a data table or a set of data tables and sub-tables, depending on the complexity of the database. A record from this database, showing the information for a single customer, may appear to the user issuing a query to the database as a collection of selected fields in a pre-determined format for that query, perhaps as a mailing label with only the name and address, or perhaps as a complete dossier with the contact information and a record of past transactions for that one customer derived from related databases. In addition to requesting a record from this database through a query, a user may ask for a report based on selected fields across many records, for instance the names of all marketing contacts within a particular state, ordered numerically by zip code and then alphabetically by last name.⁸

Databases systems tend to be highly unique and customized to support a specific task or system owner. Thus, in addition to the context typically required to understand the significance of a traditional document, the ability to fully understand the unstructured data within a database requires knowledge of data relationships, what the information represents, and how it was generated. Without this information, analyzing databases is akin to seeing a thousand-piece jigsaw puzzle without an illustration that shows the final completed puzzle. The jigsaw puzzle can be assembled, but only with great effort and with low efficiency.

B. Components of a Typical Database System

Database systems typically consist of the following elements:

Database application – a software program or programs, usually designed for a specific purpose, and usually providing a ‘higher-level’ view of the data (often through a graphical user interface) that conceals the complexity of data decomposition and data location.

8 This description of a database with its structured data should be distinguished from the term “data compilation,” introduced in the 1970 amendment to Rule 34 the Federal Rules of Civil Procedure, long before the advent of the desktop PC and off-the-shelf database software. That term was intended to encompass all of what we think of today as “electronically stored information,” and was occasionally used by courts interchangeably with the term “database,” even though the “records” in such “databases” may have included unstructured data. *See, e.g., Fauteck v. Montgomery Ward & Co.*, 91 F.R.D. 393 (N.D. Ill. 1980) (machine-readable employment records).

Database engine – the software program that stores and retrieves data at a basic level and interfaces between the applications and the database files. For example, a database engine may enforce rules pertaining to data such as only allowing storage of numbers in a telephone number field or ensuring that all invoices pertain to a customer.

Set of structured tables or files – These contain the substantive data, often in a vendor-specific format.

Confusion can arise when parties use the same terminology to describe all three components of a database system.

The individual parts of a database system may themselves be composed of multiple parts. A *database engine* may be composed of multiple software programs that collectively provide core database functionality in a given hardware and operating system environment. The *database application* may be composed of tens – or hundreds – of individual programs. The *database storage file* that typically contains the information relevant to a specific legal dispute may be a single file, but more commonly, it is composed of multiple separate data storage files in multiple locations. Large storage systems may be composed of hundreds of separate data files.

C. Assessing Relevance for Databases and Database Records

For the reasons given in section A. and B. above, the legal team often will require the assistance of individuals with technical and business expertise in order to assess what information within a database system is responsive to a particular matter. Although a database system may contain relevant, even critical, information, it also may contain information that is irrelevant or only tangentially related to the issues in a particular case. For example, the financial accounting system used by a large company may contain thousands of different data tables and tens of thousands of data fields. In most cases, however, only the substantive information contained in a small number of tables or fields will contain information of direct relevance to a legal dispute, unless the dispute relates specifically to the design or performance of the system. Thus, working successfully with a database system requires understanding how information is organized within a database and the relevance of the various fields to the issues.

To identify the data that might be relevant in a particular matter, the legal team must understand the core issues of the case, the facts that might prove or disprove liability, and the factors that might be useful in establishing or refuting damages. Different types of cases will require different types of information and will make use of database information in different ways.

D. Preservation of Databases

A party is obligated to take reasonable steps to prevent the deletion or modification of information in its possession, custody or control that it knows or reasonably should know is relevant to pending or reasonably anticipated litigation. This obligation applies to databases, but differs from preservation of unstructured ESI in a number of important ways. Preservation of information contained in databases usually requires expertise of database system or application administrators. For certain information

in databases that is not overwritten (and is essentially aggregated) it is reasonable to preserve the data “in place,” but for other dynamic data that is not stable it may not be technically possible to preserve the data “in place.” For instance, if the data is volatile (subject to being programmatically changed or deleted) or if the database system or application has enforced retention periods that for technical reasons cannot be readily suspended or interrupted, then it may be advisable to copy the specific responsive information to a separate secure location in a manner that protects that responsive data. Because of the expense of production, restoring, and interpreting backups from tape or disk, preservation by means of backups should only be used in situations where there is no other reasonable means of preservation. One thing that is consistent across databases and unstructured data is that responding parties are only obligated to take steps to preserve the information that is actually relevant to the matter and not all data within the database or in the data source.

E. Collecting and Producing Database Information

Differences in ways that database information and individual documents are organized also require different approaches and tools in the traditional discovery tasks of collection, review, and production. Unlike loose documents, database information does not fit neatly into standard document collection protocols. It is in the interests of both requesting and responding parties to avoid over-production of information. Other than situations where a large portion of a given database is responsive, it may be best practice to collect that responsive data by saving a copy of a subset of the database information to a separate location, such as a specifically-designed table, a separate database, or a text delimited file by means of a query or report. In some cases, a pre-existing (‘canned’) query or report may exist that can be used for this purpose. In other cases, a custom-created query or report will need to be used.⁹

Assuming that one can create a separate copy of a subset of relevant information from the database, the format by which this will be produced should be considered. Unlike text delimited files, a given database format will often not be readable by other software. Therefore, both parties should communicate early about the format for production so that the ESI is reasonably usable by the receiving party in accordance with Rule 34.

These uniqueness and customization issues preclude the use of generic ESI collection tools to capture relevant information within a database. Consequently, the process for understanding and retrieving the data from databases can require significant “hands-on” involvement by the database managers as well as database users to educate the legal team about the contents and structure of the database in question. This process is often matter-specific and potentially labor-intensive.

Certain specific types of contextual information are commonly requested and produced from databases. These include:

- *Field names*, which may or may not help the requesting party understand the contents of each field. Note that field names and field contents may not necessarily be related, as in databases that have been in use for some time or whose primary design objectives have changed.

⁹ Note that if a text delimited file is produced and the format does not have column headings, then it is also generally necessary to produce metadata to explain the fields in the text file.

- *Field values and codes*, which define any abbreviations stored in data fields. Field codes, whether abbreviated or not, may require further context to convey their meaning to a requesting party. For example, the code “SG” that is stored in the Product Category field might require both translation to “Sporting Goods” and a further description of what this term encompasses within the organization. Field value translations and/or associated lookup tables may be critical to understanding accurately the content of the data file, and a responding party should provide this additional information if necessary.
- *Input constraints*, that describe the allowable and/or expected values in a field. Common examples of field input constraints include numeric-only limits, state code abbreviations, and ZIP code validation. Understanding field constraints can explain why the data has been standardized in a specific way. Conversely, knowledge of these input constraints can make it easy to check a data production for errors; abnormal field values in the production may indicate that there were errors in process used to extract and prepare the data for production.
- *Auto-filled fields*, such as username or time stamps, are populated automatically by the system and without human intervention. These fields may be valuable validation tools in the ordinary course of business, as they are unlikely to contain human data entry errors, and they may have similar value in authenticating database information for possible evidentiary use. A requesting party may find it valuable to request the identification of these fields, along with the rules or programming logic used to populate them.

Information contained in databases may be the best source for establishing certain facts in a legal dispute. Information stored in this format also may be useful, if not essential, for analyses such as sorting, calculating, and linking to answer quantitative questions presented in a case. In contrast, documents such as individual email messages and free-form electronic word processing and presentation documents are not easily calculated or sorted based on their content, though they may better answer certain qualitative (as opposed to quantitative) questions than database information. Information extracted from databases is often used by accounting or economics experts on behalf of litigants, who use the quantitative conclusions of these analyses to support their legal positions.

F. Potential Use of Database Information by a Requesting Party

An important consideration in how database information should be requested and produced in civil litigation or regulatory discovery is the manner in which the requesting party intends to use the information. Without such mutual understanding, databases and database information may be produced in ways – even electronic, machine-readable formats – that are not suitable for the requesting party’s needs. A requesting party may use structured ESI in a variety of ways, including, but not limited to: (1) reviewing specific historical transactions and records; (2) developing an archive of information that can be queried as might have been done in the ordinary course of business; or (3) developing new analyses of the information that are based on a current, not historical, understanding of the data. The anticipated use of the data will drive the discussion regarding the most appropriate production format for structured ESI from a database system.

Reviewing historical information typically requires the simplest production format of these three potential uses. If the parties are interested in discrete transactions or events, a simple query or review of the data to isolate relevant records may be sufficient. A simple example of this use would be querying a database for information regarding a specific invoice. Depending on the volume of information required for this use, database information can be produced in a number of different production formats, possibly even those that do not preserve the fielded nature of the information. Simple “canned” reports displaying the requested information may be adequate, and such reports sometimes can be exported into standard electronic formatted files, such as Microsoft Excel, or Comma Separated Value (.CSV).

However, developing an archive of relevant information that can be queried as might have been done in the ordinary course of business may require a more elaborate production format. For example, if the dispute involves all invoices and other interactions with a particular customer, relevant information may include a large volume of invoices and other accounting information, as well as standard reports that were generated or used by key players in the dispute as the basis for decisions involving that customer. Sometimes, the requesting party also may want to replicate standard reports that were used by the producing party, but with altered parameters, such as generating reports based on quarterly instead of annual data.

For purposes of deciding a production format, one key consideration is whether the requesting party will need to generate various alternative reports using a variety of search parameters. If so, then it is likely that the requesting party will need to receive not only the source data, but also a means to edit the “canned” reports, or create new reports. However, when the relevant information is contained in only a few set reports, the producing party may be in the best position to generate and produce the specific reports, to the requesting party.

The need of a requesting party to develop new queries and reports to analyze the data from an existing, and particularly legacy, database system can raise the greatest challenges to identifying and implementing a useful production format for database information. For example, when a requesting party has a legitimate need to develop an independent analysis or show the significance of viewing the data in a certain way, responsive data must be provided in a format that supports the legitimate intended use. As such, the requesting party must make reasonable efforts to work with the responding party to ensure that structured ESI extracted from a database is produced in an appropriate reasonably usable format. This can be a complicated process for the producing party, particularly if the requesting party seeks the underlying data in a format in which it ordinarily has not been stored. When such situations arise, the parties should consider the scope of the request and the cost and effort required to collect and produce the information from the database in a reasonably usable format.¹⁰

The data analysis undertaken by a requesting party can range from simple data accumulations, such as total sales in a given time period, to complex time trending that reveals specific patterns in the data. Often, the requesting party will need to create custom reports or new tables to support these analyses. To ensure the accuracy of the underlying source data on which these analyses are based, at times it may be necessary to produce operational manuals, schematics, or other ancillary documentation that is required for the requesting party to correctly assemble the data.

¹⁰ See, Database Principle 2, Accessibility and Proportionality.

Creating new analyses of information contained in a database often expands a discovery request beyond the immediate fields that contain the substantive information at issue. For example, a call center application may have components that help manage the workflow between the agents. This may include external logs that track who participated in a particular call, how the call was processed, and its ultimate disposition. Even if the responding party does not routinely look at all of this stored information, if there is a question as to how the responding party managed its calls, then a requesting party may reasonably want to analyze this data, including the internal system fields that are not visible to the user that tie these disparate data elements together. Therefore, it is critical for the parties to confer as to the scope and format of the information to be produced.

A final consideration with respect to the requesting party's need to perform new analyses on the structured ESI is the extent to which the requested information can be introduced as substantive evidence in court. While the traditional approach for introducing this type of electronic evidence is through a testifying expert, some testifying experts may not be qualified to manipulate the underlying data to create the analysis that may form a partial basis for their conclusions. Certain experts may instead work with one or more technicians who serve as the interface between the data and the testifying expert. At this time, there are no standard practices with respect to these data technicians, and it is unclear to what extent their activities must be validated or whether they themselves must be available to testify as fact or expert witnesses to meet the evidentiary requirements. Further, such data processing has at times introduced questions regarding the accuracy and admissibility of analyses, even though they are based on the original data produced in discovery by an opponent.

G. Locating Specific Database Information through Queries

Counsel should adequately communicate with the information technologists, database users, or other client representatives responsible for the database systems to determine the most efficient way to locate the responsive data. Those who are responsible for actually identifying relevant database information may need to rely on search tools, particularly for ESI within a larger database or database system. Three basic types of tools are available for this task: (1) built-in search functions relying upon an internal database index; (2) search functions that search database content in real-time (non-indexed) searches; and (3) third-party tools that develop their own indices or search existing data tables using alternate search algorithms. However, it should be noted that the Information Technology (IT) departments in many large organizations require that such third-party tools be comprehensively tested before installation or use to ensure that data integrity and operational functionality are not impaired. In such situations, the testing protocols can be quite rigorous and time consuming, thus potentially affecting the practicality of this third option.

Database indices¹¹ can be used to speed up queries against database data. Because database indices typically reference only a subset of the data fields that exist within a database, parties may need to assess the value of using additional technology to conduct broader searches that access more or additional information within a database. However, such "database-crawling" tools can significantly impact the speed at which a database processes transactions. In considering whether such supplemental measures are required, the parties should weigh the likelihood that the search will provide useful additional

¹¹ *The Sedona Glossary* defines index as: "Database fields used to categorize and organize records. Often user-defined, these fields can be used for searching for and retrieving records."

information against the burden that this approach would place on the responding party, both in terms of litigation costs and potential business disruption. This analysis can be very fact-specific, and requires that the parties engage in an open and well-informed dialogue.¹²

H. Databases and Database Information in a Third Party's Custody or Control

It is common for companies to outsource some or even all of their IT functions to third parties – including the storage and management of database information. For example, many companies outsource their payroll function to another company that maintains some, if not all, of the detailed information regarding payroll on their databases and systems. In certain situations, information managed and maintained by these third parties could become relevant in a legal dispute and fall under a legal hold. In addition, while the substantive data sought by a requesting party may be deemed to be within the responding party's "possession, custody, and control," there may be ancillary data or metadata necessary for full understanding of the substantive data. Such information, like field structures or metadata, may be in the hands of a vendor or service provider, requiring a subpoena under Rule 45 to obtain. While the situation of potentially relevant data being stored at a third party location outside the possession, custody, or control of a litigant is not new or even limited to ESI, discovery of database information stored in a third-party repository can involve a complex mix of competing rights and obligations that may require court intervention to resolve.

When data is housed by third parties (e.g., "cloud computing"), it can complicate the legal and technical issues related to data preservation and production. These issues are beyond the scope of this Commentary, but some of the important issues to keep in mind are:

- Whether a party can legally obtain requested database information from the third party and the costs involved, which may be governed by the terms of a service contract.
- The extent to which the requested data may be co-located with data of other non-parties, and the difficulty of extracting only the requested data.
- The extent to which proprietary information, software, or equipment of the third party is required to understand or use the requested data.
- The extent to which the integrity or management of the data by the third party is itself a relevant issue in the litigation.
- Whether in any particular litigation, it is more appropriate or efficient to request an opposing party to produce the data under Rule 34, or request a third party to produce the data under Rule 45.

12 See, e.g., *Soto v. Genentech, Inc.*, 2008 WL 4621832 (S.D. Fla. Oct. 17, 2008) (producing party failed to provide sufficiently detailed information to support its burdensomeness argument as to the time and effort required to compile certain relevant information stored in databases). See also, *FDIC v. Brudnicki*, No. 5:12-cv-00398-RS-GRJ, 2013 WL 2948098 (N.D. FL, Panama City Division, June 14, 2013) (rejecting the argument that proposed database search protocol requiring parties to collaborate in creating search terms was unduly burdensome and permitting modest cost-shifting consistent with traditional paper cost-shifting).

II. APPLICATION OF THE EXISTING SEDONA PRINCIPLES TO DATABASES & DATABASE INFORMATION

Since 2003, *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*¹³ has provided guidance to the legal community for the preservation and production of all forms of ESI, including databases. In Section III of this Commentary, we propose six new Database Principles that specifically address the issues associated with databases and database information. However, discussion of how the existing Sedona Principles (Second Edition, June 2007), particularly Principles 3, 5, 6 and 12, apply to the discovery of databases and database information is instructive.

A. Sedona Principle 3: The Early “Meet and Confer”

Parties should confer early in discovery regarding the preservation and production of electronically stored information when these matters are at issue in the litigation and seek to agree on the scope of each party’s rights and responsibilities.

Sedona Principle 3 is especially applicable in the context of database discovery because of the complicated technical and logistical questions raised by the storage of information in database systems. Database discovery may entail some of the most expensive and complex discovery in a litigation matter, and meaningful conversations between the parties early in the litigation can substantially reduce confusion and waste of resources. It may be in the best interest of the parties to meet and confer regarding the specific fields that contain relevant information, and the specific exports and production format.

By addressing issues related to the preservation and production of information stored in databases as early as possible, parties can resolve easier questions and make progress on resolving more difficult ones. Sharing technical information also may benefit a responding party by educating the requesting party as to what information exists. Such early disclosure can help a responding party avoid wasting resources looking for data that does not exist or that the requesting party does not actually intend to use. Similarly, early discussion may identify specific cost or burden points that can be resolved relatively easily. For example, an ongoing preservation¹⁴ would involve continually preserving every change to a dynamic data field, can be time consuming, expensive, and may not be practical in certain database systems. Advised of this, a requesting party may find that it needs only a single snapshot of that information, sparing the responding party unnecessary preservation costs.

1. Redactions, Omitted Data Fields, and the Inadvertent Production of Privileged and Other Protected Data

While a database that logs the use of electronic key cards for entrance into a building is unlikely to contain any attorney-client communications or work-product materials, some databases may contain granular information that requires special protection. For example, a database may contain personally identifying information, such as Social Security numbers, of the people using the key codes. Similarly, a database system that is used to manage a work flow for creating and publishing promotional material may store

13 *The Sedona Principles Addressing Electronic Document Production*, (2d ed. 2007) <https://thesedonaconference.org/download-pub/81>.

14 Ongoing preservation is not only of historical information that pertains to certain conditions, but also of any new information coming into the system pertaining to those same conditions.

comments from the in-house or retained legal counsel regarding the materials that fall under the attorney-client privilege. Such privileged notations may be placed in discrete “attorney notes” fields that could be isolated, or they could be mixed with non-privileged comments in free-text data fields.¹⁵

Early conversations between counsel regarding the existence of protected database information and how that database information should be treated can reduce costs and burden on both sides. For example, both sides may agree that the responding party need not disclose its employees’ or third-parties’ Social Security numbers, thus sparing the requesting party the need to set up complicated protective structures to comply with privacy laws or regulations. However, that may not always be possible. Using the earlier example of privileged communications that may be mixed with other free-form notes, the requesting party still may seek production of this field, with any privileged communications redacted and logged. Under such circumstances, the responding party may be required to budget for and execute a review of the database content, creation of a database-specific privilege log, and development of a protocol that clearly identifies the redaction of this content without otherwise disturbing the integrity of the rest of the data being produced.

It is good practice to discuss the topic of redaction early in discovery in general, and even more so with redaction of database information. Redaction of database information can take two basic forms: (1) not producing a field of information; and (2) overwriting some or all information in a data field so that the requesting party can see that information had been stored in the field. Early discussion can yield agreement on the type of redaction applied to protected information, such as replacing text with strings of uncommon characters (e.g., “&” or “@”) to make it easy to find redacted information at any point. Deferring this conversation until later in the discovery process complicates and adds expense to the production of database information, as information may have to be treated more than once to meet the protocol that is ultimately negotiated.

Another database production issue that benefits from early conversation is the treatment of information that is inadvertently disclosed. Because database information is not well suited for inclusion into most, possibly not any, document review platforms, this information may not be scrutinized as closely as the discrete electronic files and email messages that make up the bulk of most ESI productions. As a result, the risk of inadvertently producing protected personally identifiable information may be higher in productions of database information than in production of other forms of ESI. Accordingly, parties are well advised to discuss protocols and consequences of producing or encountering inadvertently produced database information, including stipulation to an appropriate protective order. *See, The Sedona Conference Commentary on the Protection of Privileged ESI.*¹⁶

2. Use and Role of Consultants and Technology Partners

Discovery of database information differs in many respects from discovery of email and file-based ESI, and data collection and review of databases are the two phases of the

¹⁵ *See, e.g., Chen-Oster v. Goldman Sachs & Co.*, No. 10 Civ. 6950(AT)(JCF 2013 WL 3009489 (June 18 2013) (finding information in data fields are communications subject to attorney-client privilege and denying motion to compel until plaintiff could offer evidence of a waiver).

¹⁶ *The Sedona Conference Commentary on the Protection of Privileged ESI*, (2014) <https://thesedonaconference.org/download-pub/3936>.

discovery lifecycle that vary most dramatically. The technical and logistical nuances in producing and receiving information extracted from databases create many opportunities for errors in the process. Thus, responding parties and their counsel may wish to use consultants and other technology partners to assist in preserving, extracting, analyzing, and producing data from databases. Likewise, requesting parties may want to employ subject matter experts to help analyze and understand the database information received in discovery. Involving these consultants early in the litigation, at the meet-and-confer stage if not before, can save all parties significant time and money, and help prevent miscommunication and duplication of effort.

It must be noted that not all e-discovery consultants have the requisite understanding of the technical aspects of database discovery, and parties should be careful to ensure any potential consultants have the actual expertise to address and resolve the database discovery issues present for the particular situation. For example, consultants and technology partners used by the responding party should understand that standard forensic collection practices may not be applicable to large enterprise databases and that separate verification and validation procedures may be required for extracted data. Consultants for receiving parties should be familiar with ways to review extracted database information. Analyzing email messages and discrete electronic files typically involves a team (sometimes a large team) of reviewers and takes place through a document review platform. Such review and analytical tools, however, are a poor fit for the matrices of information found in tables of extracted database information. Instead, review of this information may require technically sophisticated analysts to query the data and extract the meaning of its aggregated information.

Few, if any, industry standards exist to measure the competence of database discovery experts and consultants. As always, when considering a potential technology partner, parties should consider the qualifications of the partner, the cost, and the defensibility of the solutions and processes that these experts suggest for the legal dispute.

3. Impact of Remote Jurisdiction and Location

While beyond the scope of these Principles, it is important to understand that large enterprise-wide databases may pull data from multiple physical locations, including data stored outside the United States. Moreover, some U.S. companies make substantial use of databases that are stored entirely on computers outside the United States and are available only through remote access. Either of these situations may require parties to consider not only their respective needs in the immediate legal dispute, but also whether laws of foreign jurisdictions will complicate or even bar the use of database information outside the jurisdiction where the information is stored. Parties should discuss these issues early on to understand the impact of these logistical and legal limitations. Additional guidance may be found in *The Sedona Conference Framework for Analysis of Cross-Border Discovery Conflicts*,¹⁷ published by The Sedona Conference Working Group 6 on International Electronic Information Management, Discovery and Disclosure (WG6).

¹⁷ See, *The Sedona Conference Framework for Analysis of Cross-Border Discovery Conflicts*, (2008) <https://thesedonaconference.org/download-pub/67>.

B. Sedona Principle 5: Duty of Preservation

The obligation to preserve electronically stored information requires reasonable and good faith efforts to retain information that may be relevant to pending or threatened litigation. However, it is unreasonable to expect parties to take every conceivable step to preserve all potentially relevant electronically stored information.¹⁸

Preservation of databases and database information can take place in a number of ways; the database structure and nature of the data it holds likely will suggest an appropriate procedure to ensure that potentially relevant data is not inadvertently altered or destroyed. The mere fact that a database contains some relevant information does not necessarily mean all information in the entire database must be placed under a legal hold. Database analysis typically starts with the most granular or atomic level possible – individual data fields – and uses relevance to guide the determination of whether information in that field should be preserved pursuant to a legal hold.

When preservation involves saving the results of a custom query or report outside the database, the specific query or report which was used to create the results also should be preserved. If preservation is done ‘in place,’ it is good practice to save both the query and report that was run, as well as a copy of the produced data.

1. Burden of Preservation

The burden of preserving a database may be relatively modest if the system maintains all information that has been entered into it – i.e., the repository serves as a permanent archive as well as a source of current information. In such cases, while the exact state of the database may change over time due to the addition of new records and information, there is less of a risk that information that existed at the time that a preservation obligation arose will be lost. Similarly, if a company’s retention policy and practice is to permanently retain in the database the ESI that is relevant to the claims and defenses in the case, preservation in place may be an acceptable way to meet the preservation obligation.

On the other hand, preserving database information may be more complicated when it is stored in a system that purges database records and information on a routine basis.¹⁹ Just as some email servers may retain messages for short periods of time before automatically deleting them, some transactional databases also remove records after their information has become dated or is no longer required for ongoing operations. One approach taken to preserve such transactional information is to retain archival or disaster recovery media for the systems that capture and process the transactions. Unfortunately, this broad preservation approach includes not only potentially relevant data, but also all of the data on the system. In addition, storing historical data in this format can strain IT resources

¹⁸ For additional guidance, see also *The Sedona Conference Commentary on Legal Holds: The Trigger & The Process*, 11 SEDONA CONF. J. 265 (2010); *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013).

¹⁹ JAY E. GREINIG & WILLIAM C. GLEISNER, *E-DISCOVERY & DIGITAL EVIDENCE* § 7.18 (2008) (finding that the scope and format of preservation as it relates to structured data is not straightforward, as the data is generally composed not only of the individual pieces of data, but also a method of interconnecting such data); *Paul v. USIS Commercial Servs., Inc.*, 2007 WL 2727222, slip op. at *1 (D. Colo. Sept. 17, 2007) (court declines to shift \$292,000 in preservation costs after the parties failed to agree to narrow the scope of database discovery); see also Thomas Y. Allman, *Managing Preservation Obligations After the 2006 Federal E-Discovery Amendments*, 13 RICH. J. L. & TECH. 9, at § 46 (Spring 2007) (When data is automatically and frequently overwritten, “preservation obligations can be difficult or impossible to execute.”).

and disrupt business operations, as well as lead to substantial downstream costs when the database must be recreated as part of the process of restoring information from archival or disaster recovery media.²⁰

In situations where a database lacks a permanent archival function or where there is no reasonable way to interrupt the usual purge or deletion cycles in order to support data preservation during the expected duration of the legal hold, preserving the relevant information stored within the database may require exporting a copy of some or all of the information to a more permanent storage medium. Tools that can accomplish this task include data export functions (either to static data tables or to an alternate database platform), special backups of the database (or of an appropriate portion of it), or by using built-in or third-party report writing functionality to identify, organize, and output the relevant information.²¹

2. Inventory and Default Retention Periods

Because of their complexity, databases often will require additional expertise beyond that of a legal team familiar with working with other sources of ESI, such as email messages and discrete files. In addition to understanding their databases and the information stored in them, parties should also be familiar with how databases may interact with one another and whether the information in the databases is permanent or transient – i.e., is deleted or purged from the database after a set period of time or when specific conditions are met.

Many databases are subject to update and modification as part of the normal course of business. In addition, practical business considerations may prevent a party from locking down data contained in a critical database. In such cases, it is critical that the party develop an alternative way to preserve the relevant ESI. For example, if the prices or product offerings of an online retailer is relevant to the claims or defenses of a case, and preventing changes to the underlying pricing and product databases that control the products available to customers would impose an undue burden on the retailer, the party could preserve the relevant ESI outside the database in the manner described in B.1. above.²² The retailer should, however, take proactive steps to preserve such data if it becomes reasonably apparent that time-sensitive information is likely to become relevant to a legal dispute. Failure to take appropriate proactive steps has led to sanctions or adverse inference instructions when potentially relevant data has been lost because a party's normal business practices for maintaining dynamic data sources led to the destruction of potentially relevant database information after a legal hold obligation accrued.²³ In such cases, responsive data can be preserved outside the database in the manner described in Section B.

20 The Federal Rules Advisory Committee noted in 2005 that “many database programs automatically create, discard, or update information” ... and “that suspending or interrupting these features can be prohibitively expensive and burdensome.” *Id.* (internal citations omitted).

21 13 RICH. J.L. & TECH. 9, *supra*, at § 48 (proposing to battle the problem of preserving continuously changing data in on a database by running and recording queries of the database at certain periods of time).

22 See, e.g., *Playboy v. Welles*, 60 F. Supp. 2d 1050 (S.D. Cal. 1999) (discussing factors to consider before ordering shutdown of producing party's online business to harvest potentially responsive ESI).

23 See, e.g., *Linnen v. A.H. Robins Co.*, 1999 WL 462015 (Mass. Super. 1999) (adverse inference jury instruction appropriate where responding party violated ex-parte order to preserve back-up tapes).

C. Sedona Principle 6: Responsibilities of Responding Parties

Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.

1. Parties Must Understand Important Database Characteristics

At a minimum, parties participating in the discovery of database information should familiarize themselves with a number of basic database attributes so that they have adequate knowledge and understanding to develop reasonable procedures for preserving and producing information from these repositories.

- **Functional Purpose.** What is the purpose of the database system? A database may have field names that appear to indicate relevant information, but the actual information stored in the system may be completely different and irrelevant. Accounting systems, payroll, sales, and operations systems are database systems commonly found in many organizations. They may be critical to the ongoing company operations. However, some or even all of these systems may not contain relevant information. Understanding how data is used will help determine whether or not the database in which it resides should be subject to a litigation hold.
- **System/Business Owners.** Who are the primary users of a database? Who are the administrators who maintain the “plumbing” of a database? These two groups, which may or may not overlap, together comprise the witness pool most knowledgeable about these systems. Database administrators/managers generally have the greatest knowledge of which users have access to the data and which users can add or modify information. Database users, on the other hand, can provide critical information about the nature and value of the information in the database that will identify whether the database is likely to be relevant. These users can provide invaluable substantive information, such as formatting inconsistencies, data anomalies (e.g., when a data field becomes used in a new way and old information is not the same as new information entered into the same field), and other functional limitations.
- **Location.** Parties should know the physical location of its databases and understand how the data is managed. Because many databases are located in remote server farms (e.g., co-location facilities) or even in different countries, it is possible that the law of more than one jurisdiction may apply to any database discovery that must take place. Database systems also may be managed by third-party vendors whose proprietary database management procedures are not necessarily known to, much less legally under the custody and control of, a party.
- **Reports.** Existing report templates or “canned reports” are a valuable and low burden method for identifying and potentially producing database information. Canned reports are particularly helpful when only a subset

of the information in a database is potentially relevant. Knowing what reports are available will help a party better understand the burden of complying with database discovery requests. For example, it may be possible to provide a requesting party with 80% of the database-stored information it seeks through a canned report, with extraction of the remaining information requiring a much greater effort. When presented with this information, the requesting party may defer the remainder of its request until it has a better sense of the actual relevance of this information to the legal dispute. Canned reports themselves can often be saved into database tables, providing a requesting party with validated, reliable information that can be used as produced or as raw data for further analysis.²⁴

- **Archival, Retention and Disaster Recovery Policies.** Database systems frequently archive historical data that has exceeded its useful life and has no further business purpose. For example, online banking records often fall into this category; transaction records may be available for a discrete period of time before being archived and purged from the active database. Data that has been archived may still be accessible if required, although the burden of retrieving it is notably higher than when it was active data within the database. It is critical that parties to a potential suit know the extent to which database information is archived – and the schedules by which active and archived information is ultimately purged from a database system.
 - **Legacy Systems.** In an infrastructure-upgrade project, it may be less expensive for an organization to start fresh with a new database system than to transfer all existing information from an old system. In such cases, the “old” legacy database may be maintained or archived in case its historical information is ever required. Legacy database systems are frequently associated with accounting or operations systems that were replaced, rather than upgraded. Orphaned legacy systems – databases or systems with no identifiable users, custodians, or technical support – also are common in merger or acquisition situations, when the corporate information of one entity is no longer in active use. A party should be able to identify what, if any, relevant legacy database systems exist within its organization, as well as whether any relevant information in these systems was ported to a newer, more readily available format.
2. **The Responding Party Ordinarily Should Determine the Best and Most Reasonable Way to Identify, Extract, and Produce Relevant Data from Databases**

A responding party, with the advice of its counsel, is responsible for determining a reasonable method for identifying, preserving, extracting, and producing relevant data from

24 Producing reports from databases in lieu of production of the database itself is supported by Fed. R. Civ. P. 34(b)(2)(C)(iii): “A party need not produce the same electronically stored information in more than one form.” But at least one court has held otherwise. *Margel v. E.G.L., Gem Lab Ltd.*, 2008 WL 2224288, at *5 (S.D.N.Y. May 29, 2008) (“[I]t appears that EGL-USA’s only objection is that the database is redundant of the information that has already been produced. I do not find that objection to be persuasive in light of the fact that information maintained in an electronic database is necessarily in a form that is not identical to a report prepared on the basis of that data and should, therefore, ordinarily be produced.”). The court in *Margel* did not cite Rule 34 for this proposition, and instead cited a case that ordered a party “to produce paper and electronic copies of same documents.”

databases.²⁵ However, just as a driver of a car may need a mechanic to help understand how the automobile's engine or on-board computer works, a party may require additional expertise to develop adequate procedures to identify and produce database information. Normally, such expertise, whether through consultants, IT professionals, or other specialists, serves as an adjunct to the responding party's legal team. In highly disputed situations, however, courts may choose a neutral third party, such as a special master, to assist with this process.²⁶

3. Parties Must Consider the Database as It Is, Not as It Could Be

Databases may be in service for extended periods of time, evolving with the needs of the organizations that created them. However, older systems may be unwieldy or inefficient when compared to current or newer database applications and installations. This can lead to frustration (by all parties) with the functionality of a given database, and claims by a responding party that certain requests for information stored in a database are unduly burdensome. Requesting parties have challenged such claims of undue burden, arguing that a responding party may not rely upon idiosyncrasies and limitations in its systems to establish burden; parties may not "hide" behind a unique and burdensome data management system which they created. However, absent evidence that a party has purposefully designed its data systems to thwart discovery, such challenges are not supported by Fed. R. Civ. P. 26(b)(2)(C)(iii) and its state analogs as those rules implicitly hold that the requesting party finds the producing party's database system as it is.

A number of courts have held that absent a statutory requirement to maintain data in a specific manner or in the absence of a specific preservation obligation, a company may maintain its corporate information in any manner it chooses, so long as its system is not intentionally designed to frustrate discovery.²⁷ As a consequence, a requesting party finds a producing party and its IT systems as they are and not as they wish them to be.²⁸

This lack of explicit legal obligation does not mean that an organization should not consider litigation discovery issues and potential costs when choosing or implementing a new database. However, the organization is not required to design or implement its databases around the potential for litigation. Virtually all databases include some design compromises after balancing competing business and legal needs. Ensuring that the database can conduct core-business functions in the ordinary course of business typically is a higher priority than ensuring that the database has capabilities for the identification, collection, and production of data that is potentially relevant and responsive to litigation. Such design decisions are appropriate, as long as they are not made to frustrate legitimate discovery.

Not all courts have held that self-imposed idiosyncrasies of a litigant's information management systems that make it challenging or costly to extract information in response to

25 See *In re Ford Motor Co.*, 345 F.3d 1315 (11th Cir. 2003) (stating that the responding party's choice to review database and produce only those relevant portions was adequate discovery response absent specific evidence to the contrary).

26 *Maggette v. BL Dev. Corp.*, 2010 WL 3522798 (N.D. Miss. Sept. 2, 2010) (inability of a party to retrieve relevant information from one or more of its databases over the course of five years, required the appointment of a special master).

27 See *Arthur Andersen LLP v. U.S.*, 544 U.S. 696, 704 (2005) (endorsing business practice of routine records destruction).

28 *Jones v. Goord*, 2002 WL 1007614 (S.D.N.Y. May 16, 2002) *claim dismissed*, *Jones v. Goord*, 435 F. Supp. 2d 221, 266 (S.D.N.Y. 2006). The court described the interconnected and interrelatedness of the data as follows:

[T]he databases in question are not simply collections of lists or numbers that can be easily extracted and correlated with other numbers; rather, each of the requested databases has been constructed to support the interactions of hundreds of concurrent users rather than to support the analytical activities of a few. Consequently, the databases are integrally connected to a data system that comprises 25 separate but interdependent subsystems that each are comprised of scores of programs, tens of databases and scores of screen and report formats. There are over 3,000 programs containing a total of 1,500,000 lines of program instructions.

Goord, 2002 WL 1007614, at *10 (internal quotations and citations omitted).

discovery requests are valid grounds for limiting discovery requests due to undue burden. In this line of cases, courts have applied the general principle that a litigant ordinarily bears the costs of collecting and producing relevant discoverable evidentiary materials, even if the litigant's discovery costs are unusually high due to the way that the responding party has chosen to organize its business records.²⁹ But high costs should factor into the courts' proportionality analysis, unless the party purposely designed its data systems to thwart discovery.

When analyzing production difficulties due to limitations in a database design, underlying database engine functionality, or data integrity, parties should consider a variety of data production options to see which best meets the needs of both requesting and producing parties. For example, it may be possible to extract and produce relevant data with relatively modest burden if it is bundled with some amount of non-responsive database information. In this circumstance, particularly if the responding party produces the data as it has been kept in the ordinary course of business, such a production may satisfy the responding party's obligations, so long as the burden of extracting responsive data is roughly equal for both parties.³⁰

4. Direct Examination of Databases

Absent the parties' specific agreement, a requesting party is rarely granted permission to conduct a direct examination of a responding party's database to view or obtain information stored within it. As also noted in the commentary to Sedona Principle 6 above, most litigation discovery requests relate to a database's content, not how it operates. Allowing full access to a responding party's database makes it difficult, if not impossible, to prevent the requesting party from accessing irrelevant or privileged information; all data fields in all database records are theoretically accessible. Direct access to a proprietary database by a non-employee also may compromise the validation of the data in the database, reducing the database's reliability for both business and legal situations.

All this said, in certain civil litigation matters, responding parties have, in fact, invited requesting parties to access one or more of their database systems as an alternative to producing relevant information by exporting it or by cloning the database.³¹ Typically, the databases in these cases contain no personally identifiable information; for example, a database of manufacturing information. Typically, too, the requesting party is often supervised, either by a responding party representative or by a neutral third party. In some cases, the requesting party has agreed not to directly access the system, instead directing an employee of the responding party to enter queries and otherwise manipulate the system. Finally, the requesting party usually must sign stringent confidentiality agreements to prevent the inadvertent disclosure of any proprietary information (relevant or irrelevant) that the requesting party may see when accessing the database.

29 See *Static Control Components, Inc. v. Lexmark Int'l*, 2006 WL 897218, at *4 (E.D. Ky. Apr. 5, 2006) ("The Federal Rules do not permit Lexmark to hide behind its peculiar computer system as an excuse for not producing this information to SCC."); *In re Brand Name Prescription Drugs Antitrust Litigation*, 1995 WL 360526, at *2 (N.D. Ill. June 15, 1995) (Producing party cannot shift discovery costs to class action plaintiffs where "the costliness of the discovery procedure involved is ... a product of the defendant's record-keeping scheme over which the [plaintiffs have] no control."); see also *Dunn v. Midwestern Indemnity*, 88 F.R.D. 191, 197-98 (S.D. Ohio 1980); *Kozlowski, PPA v. Sears, Roebuck and Co.*, 73 F.R.D. 73, 76 (D. Mass. 1976).

30 It should be noted that a responding party is never obligated to produce non-relevant information. See, Section III. A. Comment 1.F. *infra*.

31 See, e.g., *OpenTV v. Liberate Tech.*, 219 F.R.D. 474, 475 (N.D. Cal. 2003) (in software patent infringement suit, responding party offers to grant requesting party access to its extensive source code database, but court orders parties to share cost of data extraction).

Direct access to a party's database systems is disfavored and has been granted over objection only in extraordinary circumstances. *In re Ford Motor Co.*³² is a rare case that discusses this issue directly in the context of database discovery. The plaintiff had requested direct access to Ford's databases to conduct queries for claims related to defective seatbelts. However, the court held that "Rule 34(a) does not grant unrestricted, direct access to a respondent's database compilations. Instead, Rule 34(a) allows a requesting party to inspect and copy the product – whether it be a document, disk, or other device – resulting from the respondent's translation of the data into a reasonably usable form."³³ The court further explained that Rule 34(a) contemplates that the responding party will search its own records directly to produce the records, not that the requesting party directly searches the data itself.³⁴ The court held that while some kind of direct access might be permissible in certain cases, this case was not one of them, because the plaintiff's request was too broad in scope and because the district court made no findings that Ford had failed to comply with discovery requests.³⁵

5. Documentation and Validation of Database Collections

When extracting data from databases for production, it is important to document, test, and validate the procedures that are used. Well-documented data collection and production procedures enable a responding party to demonstrate its good faith efforts to accurately export and produce database information. The same documentation also makes it possible to respond to any allegations of over- or under-collection of database information.

6. Features and Limitations of the Technology and Tools that can be Applied to Databases to Identify and Extract Relevant Information

Databases differ in the types of functions that are incorporated into them. For example, some databases support open-ended free-form text fields; others impose much shorter character length limitations on their data fields. All databases offer search query functionality, but some database engines support deeper search functionality than others. Still other database engines may offer powerful search features, but may index only the first several hundred characters in a data field, making standard search queries unreliable when applied to long, free-form data fields.

Responding parties have an obligation to understand the features – and shortcomings – of the database engines that power their information repositories. Understanding this technology is separate from the data content or system usage knowledge required to explain the significance of database field names or how information was entered into the structure. Indeed, different individuals within an organization typically have one, but not both, of these distinct bodies of knowledge about its databases.

Understanding the limitations of a database also requires an understanding of which external utilities – if any – can be used to add functionality to a database. For

32 *In re Ford Motor Co.*, 345 F.3d 1315, *supra*.

33 *Id.* at 1316-1317.

34 *Id.* at 1317.

35 *Id.* See also *Cummings v. General Motors Co.*, 2002 WL 32713320 (W.D. Okla. June 18, 2002); *Butler v. Kmart Corp.*, 2007 WL 2406982, at *3 (N.D. Miss. Aug. 20, 2007); but see *Qualcomm, Inc. v. Broadcom Corp.*, 2007 WL 935617 (S.D. Cal. Mar. 13, 2007) (to resolve discovery dispute over search terms applied to a proprietary Oracle database, the Court ordered the responding party to provide the requesting party access to a full version of the database, including the same search capability and client tools used by producing party engineers, along with a one-hour live training tutorial and written instructions on how to use the search tools).

example, the software that powers many enterprise-class databases may be relatively limited in the ways that it can format information into reports. Instead, these database engines allow close integration with third-party report generation tools. Because of the variety of ways that a database can store its information, however, not all reporting or other enhanced functionality tools will work with all databases or database systems.

A responding party may not be able to meet its database discovery obligations without solid knowledge of these tools and their potential application to the party's relevant databases. Without this understanding, it is difficult for a responding party to fully understand, much less articulate, the burden that a given discovery request imposes on it. Moreover, a lack of this knowledge greatly limits a party's ability to have comprehensive, frank discussions about database discovery.

D. Sedona Principle 12: Form of Production and Metadata

Absent party agreement or court order specifying the form or forms of production, production should be made in the form or forms in which the information is ordinarily maintained or in a reasonably usable form, taking into account the need to produce reasonably accessible metadata that will enable the receiving party to have the same ability to access, search, and display the information as the producing party where appropriate or necessary in light of the nature of the information and the needs of the case.

1. Mismatch of "Native Format" to Most Database Productions

Rule 34(b)(ii) and its state equivalents mandate that a responding party must produce ESI in either the form or forms in which it is ordinarily maintained (sometimes called "native format") or in a reasonably useable form or forms.³⁶ However, "native format" may not have as clear a meaning in a database context as it does for other forms of ESI.³⁷ In fact, in many cases, a truly native format production of database information is less usable to a requesting party than an alternative production format.

Database engines typically compact the information they store and index to reduce storage requirements and speed information retrieval. Each database engine uses a different proprietary format for the data files that make up the components the database uses to properly function. For example, Microsoft Access often folds all database information into a single .MDB format file. A Microsoft SQL Server database, on the other hand, is composed of several types of files, including primary files (.MDF), secondary files (.NDF), and transaction logs (.LDF). Other database engines use different structures and file types, and few, if any, can read or process information stored in a different database engine's format.

³⁶ In several instances, courts have held that databases should be produced in native format. *See, e.g., In re NVMS, LLC*, 2008 WL 4488963, at *1 (Bankr. M.D. Tenn. Mar. 21, 2008); *Covad Commc'ns Co. v. Revonet, Inc.*, 258 F.R.D. 5 (D.D.C. 2009). Compare with *Coquina Investments v. Rothstein*, N. 10-60787-civ, 2012 WL 3202273 (S.D. FL, Aug. 3, 2012) (finding that counsel should have produced a requested document in native format to preserve its original qualities but declining to award sanctions) and *In re Facebook PPC Advertising Litigation*, No. C09-03043 JF(HRL) 2011 WL 1324516, N.D. Cal, San Jose Division, Apr. 6, 2011) (ordering parties to meet and confer regarding an alternative to producing a proprietary database storage format when a PDF printout of the database did not show data fields, hence the database was not produced as it appears).

³⁷ *See, e.g., Bob Barker Co. v. Ferguson Safety Prods.*, 2006 WL 648674, at *4 (N.D. Cal. Mar. 9, 2006) (declining to order production of financial services database responsive to discovery request because "it is unclear how a party could go about producing 'a database,' which ordinarily is a dynamic collection of data that changes over time").

A true “native” production of database information provides a copy of a database that can be used only by someone possessing a licensed copy of the correct version of the database engine software. Depending on the nature and age of the original database, such a license may be difficult for a requesting party to obtain, if not practically impossible. An additional disadvantage of producing a database in its “native format” is that internal tracking may be difficult or impossible to turn off. Stated another way, this means that merely opening a database may alter some of its validation values such that the authenticity (and thus admissibility) of the database can no longer be established at the “native file” level.

While a true “native” production of database information may not be feasible or desirable, some metadata – in the generic sense of the term, “information about information” – is necessary for the production to make sense. This is a distinguishing feature of database information. As one court discussing Sedona Principle 12 put it, “while metadata may add little to one’s comprehension of a word processing document, it is often critical to understanding a database application.”³⁸ And the same court, comparing different form-of-production options, noted “one marked disadvantage of [TIFF or PDF] is that the production involves significant costs; it also does not work well for spreadsheets and databases.”³⁹

If a requesting party receives a native-file database production, the native production should be accompanied by a production of database information in the form of generic “load files” such as text delimited files that can be read by many different types of databases or other software applications. Such load files should include the fielded data that has been exported, so the requesting party can use the load files to map each information field into a database structure of its own design.

2. Use of Standard Reports to Produce Database Information

As addressed in I.E., I.F., and II.B.1, *supra*, most databases include ways for business users to view or print out multiple data fields, organized in a useful manner. The simplest database reports might present columns of information in a simple table format; more complicated reports may combine content from multiple fields, perform mathematical calculations and present them, or include graphs derived from underlying database information. Database reports may be static – that is, an unchanging view of certain data that have been selected by query, or they may be more interactive, permitting users to change the scope, focus, and perspective of the database. Generally speaking, most existing reports that are used in day-to-day business are “pre-validated,” meaning that accuracy of their data aggregation has been tested and demonstrated. Standard reports, also known as “canned” reports, should be contrasted with custom reports, where users (or database administrators) select report content based on individual or changing needs. Because these reports are created “on the fly” by database users, it is more possible for these information views to include errors, such as mismatches between field name and displayed field contents or mathematical errors.

Standard reports have both advantages and disadvantages as a production format for database information. Because these report templates already exist and have been pre-validated for accuracy, it is generally faster and cheaper to use these reports than to create custom views and information extracts. However, standardized reports may not collect all

³⁸ *Aguilar v. Immigration & Customs Enforcement Div. of U.S. Dep’t of Homeland Sec.*, 255 F.R.D. 350, 354 (S.D.N.Y. 2008).
³⁹ *Id.* at 356.

potentially responsive or relevant data in the database, and they may not produce it in the specific format that has been requested. Thus, standardized reports may be a low-burden way to make a partial production of requested database information, but they may not provide the most complete solution. If a standardized report is missing crucial data or provides the information in a way that cannot be processed using reasonable efforts by the requesting party, a different production format may be more suitable. On the other hand, if the standardized report captures all of the significant data and omits only marginally relevant information, it may be more appropriate to produce database information in a standardized report than to invest time and money into creating a custom report that provides absolutely all of the database information that has been requested.

3. Use of Fielded Tables to Produce Database Information

A common way to produce database information is through tables (i.e., rows and columns) of information, where each row represents a database record and each column represents a single data field. Most database engines, even those that do not have sophisticated reporting functionality, support exporting database information into either text delimited files or fielded tables. Similarly, many different database engines can import delimited files and separate out each field of information for subsequent analysis.

Text delimited files are closely related to, if not often virtually the same as, database “load files”; they are generically formatted sets of fielded information. Delimited files, however, may not be able to completely show the relationships found in multi-table relational databases. For example, in a banking database, a single customer may have both individual bank accounts and a shared bank account with one or more co-owners. Typically, these relationships are tracked in a multiple-table relational database, where each bank customer can be related to multiple bank accounts, and each bank account can be related to one or more customers. If this information must be consolidated in a single table, preserving these “one-to-many” relationships may require that information be repeated so that full information can be displayed in each view of the information. “De-normalizing” the data in this way (i.e., transforming it into a different format from the way in which it is stored in the ordinary course of business) is a relatively common and often acceptable data production practice, even though restoring this information into multiple relational tables to recreate the original types of relationships may not be a straightforward process, depending on the data relationships that are required.

For example, the parties could clarify whether the requesting party would prefer to see the results of a query or report that links the data elements together, or to have exports of the responsive data from separate tables and import the files into their own system in order to run their own queries.

III. THE SEDONA CONFERENCE PRINCIPLES FOR THE PRESERVATION & PRODUCTION OF DATABASES & DATABASE INFORMATION (THE “SEDONA DATABASE PRINCIPLES”)

While *The Sedona Principles* cover the preservation and production of ESI in general, and includes useful guidance for the discovery of databases and database information in particular, the complex and evolving nature of database discovery calls for a more in-depth examination of the issues that are unique to databases and the information found in them.⁴⁰ Because of the structural complexity and volume of database information, database preservation, collection and production often involves relatively greater costs and burdens than those associated with the production of unstructured media. Defining a reasonable scope of database discovery requires all parties to understand the purpose for which the information is sought, the components and respective relevance of the data at issue, the workings of the technology that stores and manipulates the data, and the processes to ensure that the data produced is what it purports to be. To that end, the following six Sedona Database Principles are intended to inform and facilitate discussions regarding assessments of relevance, potential costs and burdens, and methods for validating results that necessarily must occur between parties that are involved in database production.

1. **Scope of Discovery**

Absent a specific showing of need, a requesting party is entitled only to database fields that contain relevant information, and give context to such information, and not to the entire database in which the information resides or the underlying database application or database engine.

Comment 1.A. Database Relevance Must Be Analyzed on a Granular Level

Databases are often very large collections of disparate information. Although situations can exist when an entire database and its information are relevant to a legal dispute, often only a portion of a database is relevant.⁴¹

The process of determining which database information is relevant is performed at several levels. First, depending on the nature of the dispute, many database records will likely not contain relevant information. These normally would be excluded from production through use of search queries. Second, however, even within records that contain potentially relevant information, not all of the data fields that comprise the record may be relevant.⁴² Identifying and extracting database information in response to discovery requests requires both levels of analysis.

The process may be complicated further by the differing views available to users based upon different levels of database-security access. A database record in a database

⁴⁰ The authors also wish to call the readers' attention to *The Sedona Conference Commentary on Proportionality in Electronic Discovery* for useful guidance applicable to database discovery. See *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013).

⁴¹ See, *In re Lowe's Companies, Inc.*, 134 S.W.3d 876 (Tex. App. 2004) (granting mandamus and vacating trial court's order for retail chain to produce database for query by requesting party without any limitations as to time, location, or subject matter); *Ex parte Wal-Mart, Inc.* 809 So.2d 818 (Ala. 2001) (mandamus granted in part to restrict requesting party's access to retail chain's incident reporting database to similar incidents only). See also, *Barnes v. District of Columbia*, 289 F.R.D. 1 (D.C., Sept. 28, 2012) (granting motion to compel search algorithm because a query used to search a database and generate reports is a "writing" subject to production, but denying request to access entire database as overbroad).

⁴² See, e.g., *Bob Barker Co. v. Ferguson Safety Prods.*, 2006 WL 648674, at *4 (N.D. Cal. Mar. 2006) (declining to order production of financial services database responsive to discovery request because "it is unclear how a party could go about producing a database," which ordinarily is a dynamic collection of data that changes over time").

application that is viewed on the screen by a typical end-user generally is created from information stored on multiple data tables, and only database administrators may be able to see the raw data as it is stored in database tables and sub-tables. Unfortunately, many database discovery requests combine requests for both database records and database tables as if they were separate and mutually exclusive repositories of information. Depending on the technological sophistication of the party representatives managing this discovery, such terminology-mixing can further complicate the process of reaching consensus on the logistics of these discovery requests.

Other times, the way that database fields are organized into columns, rows, and tables may simplify conversations about the scope of production. Depending on the facts in a dispute, entire tables of database information may not be relevant and may not be required to be preserved or produced. Conversely, other data tables may contain fields of important information that require special treatment. To the extent that data is “rolled off” an active database, a database administrator may need to implement preservation measures for specific tables to reduce the risk of inadvertently destroying potentially relevant information.

Illustration i. In litigation involving a car manufacturer and the various warranties provided to consumers, plaintiffs request documents to identify the customers of certain models of cars, the cars they purchased, and the warranties they purchased. The defendant’s database that retains this relevant data also contains non-relevant information, including dealership, the salesperson, and the commission the salesperson received on selling the car. This non-relevant information is stored in the same rows and tables as the responsive, relevant information. The information in these data fields is not relevant to the dispute, and the data fields do not need to be produced. Furthermore, even though both the relevant and non-relevant information might appear in a standard view of the customer’s database record, the responding party should not be obligated to produce the non-relevant information even if the requesting party asked for “all documents related to” customers of the certain car models.

Illustration ii. In a breach of contract litigation between two companies where the amount paid by one to the other is in dispute, the defendant’s accounts-payable database could contain potentially relevant information regarding payments by the defendant to the plaintiff. However, absent a persuasive argument to the contrary, the data records (i.e., rows) regarding payments to other companies for unrelated transactions is not relevant, and need not be produced.⁴³

Illustration iii. In the same breach of contract litigation, not every data field (i.e., column) displayed in a record that contains relevant information in the accounts payable database is necessarily relevant and within the scope of discovery. For example, the “payee,” “amount,” “date,” “check number,” “approver” and “comments” data fields (and their relationship to each other) may all be relevant, but other data fields in the record may not be relevant (e.g., “unique record ID,” “tax ID,” etc. ...). *Id.*

⁴³ See *Ex parte Wal-Mart, Inc.*, 809 So.2d 818, *supra*.

Comment 1.B. *Parties Must Determine the Relevance of Individual Data Fields Within a Database*

When reviewing the relevance of data fields, parties need to carefully examine the relationship between relevant data fields and other fields (or rows, or columns, or tables), because this relationship can make otherwise irrelevant data relevant because of its link or connection to relevant information. While it is possible that a single piece of relevant data within a record or table may transform otherwise irrelevant data within the same record or table into relevant data because of their relation to each other, such a logical connection is by no means automatic.

A responding party that finds relevant information in a portion of a database should reasonably consider the entire database to determine if other portions are relevant to the dispute. A party that unilaterally examines its own databases to determine what fields are relevant or irrelevant should, as a matter of best practice, act conservatively to avoid inadvertently excluding relevant data. Generally speaking, the cost of performing this analysis a second time, plus the downstream acts of extracting and processing this information a second time, is far more than the cost of identifying, extracting, processing, or producing slightly more data during a single pass.

Analysis regarding the relevance of information contained in individual cells is not unlike that pertaining to information contained in various types of metadata. In addressing the relevance of metadata associated with various forms of ESI in *Aguilar v. Immigration & Customs Enforcement Div.*,⁴⁴ the court drew from Principle 12 of *The Sedona Principles*⁴⁵ noting that “the two ‘primary considerations’ should be the need for and the probative value of the metadata, and the extent to which the metadata will ‘enhance the functional utility of the electronic information.’” A parallel approach should be used to determine relevance of data fields (i.e., to what extent is the particular field data or its relationship to other fields essential to understanding the information sought; does such field-level data enhance the utility of the records). The *Aguilar* court noted that, “[a]s a general rule of thumb, the more interactive the application, the more important the metadata is to understanding the application’s output.”⁴⁶

If the data fields themselves are not privileged or determined to be trade secret, metadata-type database field information can be analyzed in several ways for relevance. However, in *Aguilar*, because the data was sensitive, the court suggested a quick demonstration to the plaintiffs of database functionality using dummy data stored in an otherwise identical database structure.⁴⁷ This approach could be used as an exploratory tool with a requesting party or with fact experts to gain an understanding of the overall output from the database if the parties cannot agree on the fields or cells that may be relevant to make meaningful use of the data or if the producing party lacks this level of understanding of its database systems.

Comment 1.C. *Database Relevance Is Measured by its Data, not the Application*

Under normal circumstances, a database is relevant to a legal dispute because of the database information stored within the tables or files, not the database application or

⁴⁴ *Aguilar, supra*, at 356.

⁴⁵ *Supra* note 16.

⁴⁶ *Aguilar, supra*, at 354-55, quoting *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 647 (D. Kan. 2005).

⁴⁷ *Aguilar, supra*, at 363.

database engine.⁴⁸ Unless there is a unique relationship between the database information and the mechanism that manages or displays that data (which can happen in some older or proprietary database systems), the software components of the database application and engine are unlikely to have any relevance to the discovery request, and should be considered presumptively non-responsive.

Proactively focusing database discovery requests on the data component of the system greatly simplifies the process of responding to these requests while rarely sacrificing full disclosure. Moreover, because database systems are configured for specific hardware and software environments, the effort to recreate these environments is vastly more expensive and complex than providing the data files in a format that can be loaded into whatever database systems are available to the requesting party.

Fortunately, most database information can be produced easily in a generic format that does not require a specific database engine or application to be read or analyzed. Depending on the requesting party's needs, a data file in a common form such as Microsoft Access or Excel can be produced, and allow the database information to be reasonably usable by the receiving party. Additionally, limiting database discovery to the database information which can be produced in an alternative reasonably usable tabular form obviates the need to negotiate the terms of a protective order or other limited use agreement with the non-party proprietor of the database software, cloud computing service provider, or computing platform provider.

Comment 1.D. Circumstances When a Database Application May Be Relevant

In certain circumstances the database application, structure, or even the database engine, may not only be relevant, but also essential to providing a complete response to a discovery request, for example, when the software itself either: (a) contains information relevant to the matter not otherwise stored in the database storage file; or (b) the software is the focus of one or more claims of the litigation.

Illustration i. Acme Corp. has programmed its financial system to provide a limited number of choices when categorizing financial transactions. The universe of possible choices, rather than the history of actual choices, has become an issue in litigation. Acme Corp. has been asked to produce the software application that contains the programming of these possible choices. It is clear that the database storage file will not contain this information. The parties should determine whether the production of the software is the best or only way to establish this information.

Illustration ii. It has been alleged that, for a two-year time period, Mortgage Broker Company's ("MBC") software incorrectly calculated monthly mortgage payments. MBC has been asked to produce the historical transactions, as well as the software code, that it used to calculate those transactions. It is clear that the database storage file does not contain those calculations. The parties should determine if the production of the software is the best or only way to provide information regarding the underlying algorithms used by MBC's software.

⁴⁸ See, Section I.B., *supra*, for definitions.

In some cases, it may be more valuable to understand the database application than to receive the underlying transactional data. This situation occurs most often when one set of data (“dataset A”) is acted upon by a software tool to then produce a second set of data (“dataset B”). For such discovery requests, it may be more effective to understand the software processes that transform dataset A into dataset B, rather than to simply receive dataset B, or dataset A.

Illustration iii. Franchise Food Co. tracks employee time and attendance via its point-of-sale system (“POS,” i.e., the cash registers). The POS terminals record the time that cashiers signed into and out of the system. In wage and labor litigation, it has been asked to produce all POS time entries and to produce all payroll-system records. While the presumption is that both would be produced, it may be equally sufficient or even preferable to produce the POS time entries and the software that creates the payroll system records from the POS data, rather than the static payroll-system records.

Comment 1.E. Value of Information About the Database System

In addition to disputes about the relevance of database information, or the database applications or engines themselves, requesting and responding parties often disagree about the relevance of the database system information, i.e., database’s schematics or the underlying technical information that do not concern information that is directly at the heart of the dispute, but instead seek information that may help the requesting party better understand the information that it is receiving and any limits in its accuracy or functionality. Understanding the context, origin and normal business use of ESI in a production may be helpful for the requesting party to make effective use of the data received.

Accordingly, in appropriate circumstances, a responding party may produce the database system information that is reasonably needed by the requesting party to obtain a basic requisite understanding of the structure, content and format of the data being produced, including relevant field names and values, the relational connections between data fields and tables, and the extent to which data fields are automatically populated by the system. In some circumstances, the scope of this system information may be expanded to include not just information about the specific data being produced, but also information about where the produced data originated from within a larger environment that may include multiple database servers, internal or external databases, and other related ESI. The production of such database system information might also include dependencies of the produced data on other data sources, uses of the produced data within the system or overall environment, and relationships of the produced ESI to other data within the system or the overall environment.

Illustration iv. In *Illustration iii* above, where Franchise Food Co. could have produced the POS time entries and the software that creates the payroll system records from the POS data, in lieu of the static payroll system records, Franchise Food alternatively may have been able to produce the POS time entries and, if available and reasonably accessible, background system technical information about the software that creates the payroll system records from the POS data.

Database system information may be presented in many different ways. Sometimes, tabular or graphical depictions of a complex data system, as can be found in “entity relationship” diagrams or data flow diagrams, may be both most helpful and least burdensome for a responding party to provide. Other times, it may be necessary to depose a witness with technical understanding of the system from which database information has been produced. Requesting parties should understand that there is rarely, if ever, a single, comprehensive source of the system information that they may request, and that a responding party has a burden of varying degree in collecting such information for production.

An additional consideration is that information produced from databases is rarely an exact copy of the data tables and database structure. Rather, the database information being produced is most often a subset of the sometimes substantial information that is stored in a larger database. In fact, this is often preferable.⁴⁹ Depending on the issues in the case, it may be appropriate for a requesting party to receive a description of the extraction and transformation process, including how the produced information was organized in the original database.

A final issue regarding the production of database system information is the extent to which database or system documentation is encompassed by a request for substantive information stored in a database. Organizations do not permanently retain all database system documentation they ever create, use, or reference. Absent explicit notice from opposing counsel or other extraordinary factors, a responding party should not be automatically obligated to preserve all supporting database system documentation, merely because the party has reason to believe that some ESI stored in the database may be potentially relevant to a party’s claims or defenses in a current or reasonably foreseeable litigation. Commercial documentation, in particular, is usually available from a variety of sources, including third parties. More careful analysis may be required in situations involving custom-written documentation, such as internal guides or references. For such materials, responding parties should consider the nature of the documentation, as well as the degree of unique insight that this material provides into relevant database information.

Comment 1.F. Appropriate Circumstances for Producing Additional Non-Relevant Database Information

While a responding party is not obligated to produce more data from or about a database than is relevant to the dispute, in some circumstances it may be easier, less expensive, and less burdensome to produce a larger slice of the database content or even the entire database. For example, business users of the database may have a “canned report” that compiles all requested information, plus some additional data fields. Producing this report is likely faster and much less expensive than designing a custom query and collecting the same database information through a custom data export utility. Thus, while a responding party is never obligated to produce additional irrelevant information (and may have reasons unrelated to litigation not to do so), a responding party *may* produce additional non-responsive information, so long as the responding party is not doing so for any improper purpose, such as attempting to make relevant information more difficult to extract or understand.

⁴⁹ See, Section I.E. Instead of being exact duplicates of existing data tables, the information is typically compiled from multiple tables (a “denormalized view”) and includes fewer than all fields or records stored in a given table (a “selective view”) – thus providing a variant but useful view of the data stored in the system.

2. Accessibility and Proportionality

Due to differences in the way that information is stored or programmed into a database, not all information in a database may be equally accessible, and parties should therefore apply proportionality to each component of a database to determine the marginal value of the information to the litigation and the marginal cost of collecting and producing it.

Comment 2.A. Technical Challenges to Accessibility

Information from and about databases is subject to the same rules and limitations as all other information disclosures in civil litigation, and in ordinary circumstances, information that cannot reasonably be extracted using tools that are readily-available in the normal course of business of the responding party need not be produced absent good cause and potential cost shifting.⁵⁰ Whether specific requested information within a database is “reasonably accessible” within the context of a specific legal dispute is a deeply fact-specific inquiry that must be analyzed, like questions concerning other discoverable material, under the proportionality provisions of Rule 26 and its state analogs.⁵¹

It is important to recognize the technical limitations that affect levels of accessibility, and a requesting party should never assume that all information in a database – or even all information visible to “average” database users – is equally able to be produced. Instead, once a responding party has demonstrated why certain database information or elements are more difficult to produce than others, the parties should consider whether the value of the information is worth addition burden and cost. As with other discoverable information, the parties should consider the availability of the same information in a reasonably usable form from an alternate source (e.g., printed instruction manuals, printed database reports) and whether the importance of the requested information is proportional to the additional burden or cost that would be required to extract it from the database in which it resides.⁵²

Comment 2.B. Factors for Assessing the Burden or Cost of Preserving, Collecting or Producing Database Information

A number of factors may be considered in accordance with Rule 26(b)(2)(B) and Rule 26(b)(2)(C) to determine if database information may be considered “not reasonably accessible because of undue burden or cost” or is disproportionate for purposes of

50 Rule 26(b)(2)(B) places specific limitation on the production of ESI. “A party need not provide discovery of [ESI] from sources that the party identifies as not reasonably accessible because of undue burden or cost.” *Id.* Additionally, a court on motion or on its own, must limit the scope of discovery if the discovery sought is unreasonably cumulative or duplicative, can be obtained from a more convenient source, could have been previously obtained by the party seeking the discovery or the burden or expense of the proposed discovery outweighs its likely benefit. Rule 26(b)(2). *See also The Sedona Conference Commentary on Preservation, Management and Identification of Sources of Information that are not Reasonably Accessible*, 10 SEDONA CONF. J. 281 (2009) and *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013).

51 *OpenTV v. Liberate Tech.*, 219 F.R.D. 474 (N.D. Cal. 2003) (court applies *Zubulake* factors to determine reasonable accessibility of source code database and allocation of data extraction costs); *Best Buy Stores, L.P. v. Developers Diversified Realty Corp.*, 247 F.R.D. 567 (D. Minn. 2007) (discovery of database denied when information sought was no longer in a searchable format, and database would have to be restored from original sources at a cost of at least \$124,000 with a monthly storage cost of \$27,823).

52 *See Superior Prod. P’ship d/b/a/ PBSI v. Gordon Auto Body Parts Co., Ltd.*, 2008 WL 5111184 (S.D. Ohio Dec. 2, 2008) (where plaintiff requested production of large volume of relevant documents and where deposition witness indicated that the information would not be easily retrieved from defendant’s electronic database, court recognized potential burden to defendant and ordered production of sampling of documents to allow for determination of the need to produce the rest).

preservation⁵³ or production.⁵⁴ Additionally, parties should understand that certain inherent limitations may exist impacting the production of database information.

- The extent of the ability to search on database fields.** The ability to search fields depends on the way a particular database system has been designed and the sophistication of its search engines. For example, many databases contain one or more free-form text “comments” fields that may be visible when a database record is viewed on screen. However, to optimize performance, only the more critical, defined-format fields may be indexed and searchable, with the comments fields available only once the associated record has been located. Limiting the fields that are indexed allows databases to hold large volumes of information without compromising system performance. Third party query/report generation tools are commonly used to supplement such limitations; however, these tools are not perfect solutions. It should be noted that searching or creating indices on un-indexed fields can impose a significant burden on an operational system.
- The extent to which information may be stored outside tables.** Not all information stored in a database is held in tables; it may be stored in a number of different places. For example, to facilitate speedy and consistent data entry, a database may include predefined values for certain fields, i.e., “drop down” or “lookup” tables, which may be hard-coded into the database application software itself and not stored in any searchable database fields or tables. Further, earlier entries in a lookup table may not have been retained when a table or the database itself was updated, making it functionally impossible to retrieve this system information without substantial effort and expense. Therefore, a request for production seeking all values from which an employee could have chosen while engaged in data entry might sound simple on its face, but responding to this request may be extremely difficult. Likewise, certain reports may be available within a system only as screen views and not easily converted to a printable or exportable format.
- The capability for exporting data.** Because information may be visible to a user does not necessarily mean that it is practically capable of being produced. For instance, individual-rights restrictions on viewing and exporting certain fields or the character of the fields themselves (e.g., “validation fields,” such as those that automatically capture the user ID of the person making changes) may impede or prohibit export through standard output channels. Moreover, since many databases are intended to be used as information repositories, the system may have been

53 Before even turning to the question of the burden and expense of *producing* information from a database, the party in possession of the database must weigh the burden and cost of *preserving* the database information (both its structure and its contents, the preservation of which are not always accomplished through the same means), against the likely importance of the information in resolving the issues in the case. See, Rule 26(b)(2)(C)(iii). See the discussion of Sedona Principle 5 *supra* at II.B. For additional guidance, see *The Sedona Conference Commentary on Legal Holds: The Trigger & The Process*, 11 SEDONA CONF. J. 265 (2010), *supra*; *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013), *supra*.

54 *Jones v. Goord*, 2002 WL 1007614 (S.D.N.Y. May 16, 2002), *claim dismissed*, *Jones v. Goord*, 435 F. Supp. 2d 221, 266 (S.D.N.Y. 2006) (denying plaintiffs’ motion to compel production of database maintained by the New York State Department of Correctional Services where the state made a compelling showing that the burden of production far outweighed its benefits).

designed solely with the ability for a user to add new data records or update existing records, with no functionality included for the export of records in bulk. Even extremely complex databases are often designed to be accessed by individual end users through a Graphical User Interface (GUI) through which users have the ability to view and edit a small number of records at any given time, but not the ability to export large numbers of records into a static format. To export the quantities of data often necessary to respond to civil litigation discovery requests and in a format reasonably usable to the requesting party, programmers may need to create custom tools or alternate interfaces to the database. In such conditions, the time, resources, and expense of such programming should be part of the burden analysis.

- **The reporting functionality of the database.** Some databases allow users to employ built-in or third-party utilities to search the database and format the results into a report that can be printed or exported as fielded data. Typically, an organization will create a number of standardized report “templates” from which the user can choose, and sometimes a system will allow users to craft “custom” reports. However, most reporting functions, whether template or custom, are limited in some fashion, such as in the fields that can be queried against, the number and combinations of fields that can be searched together, the volume of records that can be included in the report, or the number of characters from a given data field that can be included in the report. Additionally, certain reports may be available within a system only as screen views and not easily converted to a printable or exportable format. If a party is required to overcome these limitations in meeting their production requirements, litigation-specific reports may need to be created by programmers, requiring additional time (to create the custom reports) and resources, potentially including hard costs. Even with custom programming, it is possible that some database fields, such as system and validation fields, may not be capable of being included in a report-writing function.⁵⁵
- **The extent to which a database system is in the custody of a third-party.** In situations where a responding party has outsourced its databases systems containing responsive ESI to offsite storage solutions under the custody of a third-party referred to as “infrastructure as a service” (“IAAS”), or is using a third-party software hosting repository referred to as a “software as a service” (“SAAS”) system (e.g., Salesforce.com), the responding party may not have the direct access to the “back end” of the database that is required to implement custom programming. The parties should consider the feasibility, burden and cost of timely exporting responsive database information, and whether there is a less burdensome alternative.

⁵⁵ The reverse problem occurs when data from a legacy system or from a time before the implementation of preservation efforts exist solely in “report” format and not in the original database structure format. It may be unduly burdensome for the producing party to restore that data to the original format. Indeed, if the data is maintained only in report format in the ordinary course of business, there may be no obligation at all to convert the data into an alternate format.

- **The active or legacy status of the database.** Unlike unstructured data, where the trend generally is to consider “active” information “reasonably accessible,”⁵⁶ the fact that a database is in active use does not automatically mean that the data is easy and inexpensive to produce in litigation. Whether a database is active or in legacy status does not determine its accessibility. The same challenges in producing data from a database currently in use as in one that is no longer active (e.g., limited export functionality, poor data consistency, a limited-feature search engine), legacy databases can often pose additional challenges. For example, the software platform or operating system necessary to run a legacy database may no longer exist or can no longer be run on current hardware. Similarly, IT or business personnel who were familiar with the structure of the database may have left the organization, and it may be difficult, if not impossible, to find resources to export data or write any custom reports.

- **The availability of database system source material, if relevant.** Much of the information describing database structures and supporting hardware and software systems can be found in the end-user manuals, system documentation, written system backup procedures, training materials, and other documentation that accrues during the development or deployment of the system.
 - **Legacy Systems.** Finding documentation for legacy systems may prove much more difficult, as supporting materials (and knowledgeable employees) for systems not in active use are often no longer available after a period of time. In situations where requested supporting information for legacy systems is not available, a responding party should not be required to either create new comprehensive documentation or deconstruct the database system for the purpose of assisting the requesting party’s understanding of the system and the responsive database information.

 - **Proprietary Systems.** It also may be difficult to find comprehensive documentation for highly integrated proprietary systems, such as financial systems from SAP or Oracle, and this information may not be readily available from either the responding party or the solutions provider. Additionally, the responding party may lack actual access to certain data tables that may be a trade secret of the solutions provider, and it thus may not be possible for it to respond fully to a request for database table structure and overall organization. While the responding party should take reasonable steps to locate and produce any such relevant, but proprietary, database system information, including obtaining information from alternate sources, the courts should consider the proportionality of the burden and costs associated with licensing or otherwise locating the requested information that is not within the party’s custody and control.

⁵⁶ See *Zubulake v. UBS Warburg LLC* (“Zubulake I”) 217 F.R.D. 309, 321-22 (S.D.N.Y. 2003); *The Sedona Conference Commentary on Preservation, Management and Identification of Sources of Information that are not Reasonably Accessible* (2009), *supra*.

3. Use of Test Queries and Pilot Projects

Parties should use objective information, such as that generated from test queries, pilot projects, and interviews with persons with relevant knowledge to ascertain the burden and benefits to collect and produce information stored in databases and to reach consensus on the scope of discovery.

Comment 3.

Many disputes about the discovery of potentially relevant information stored in databases are based on deduction and inference, rather than empirical data. A requesting party may insist that certain types of information must have been stored in an opponent's database "because that's what should be there." Conversely, a responding party may estimate the burden of responding to discovery requests without ever testing whether its assumptions are accurate. Neither of these approaches is acceptable.

A better approach for establishing the benefits and burdens of producing information stored in databases is to examine objective information about the systems. To this end, a responding party may examine user manuals or any database table schematics that exist, or more incisively, use one or more queries to test how long it takes the system to return results, the effect of those queries on the system's operation, the relevance of those results to the issues in the case, and the logistics required to export this information in a format that is reasonably useful to the requesting party. Each of these objectives – the speed of the system, impact on operations, the accuracy of the query, and the data extraction – can then be fine-tuned to improve efficiency and the overall results.

Regardless of whether the responding party concludes that the information requested is "accessible," it may wish to create a test query or pilot project and share the results with the requesting party to demonstrate the steps that are being taken to respond to a discovery request and allow both sides to assess the usefulness and relevance of the exported information before incurring the cost of a full production. The test queries may identify problems with the discovery request, such as over- or under-inclusion, or the pilot project may identify issues with preparing the data for production in precisely the format requested. Sharing this information provides a common factual basis upon which the parties can re-examine the discovery requests and modify them appropriately before incurring the cost of a full production.

Illustration v. A requesting party seeks all records from a database of internal memoranda and reports that include certain key words and phrases, including the term "market." Test queries indicate that the request would flag more than two-thirds of the records as potentially relevant, even though the subject at issue is narrowly focused. A review of samples taken from the "market" query reveals that all of the sample records are, in fact, not relevant in any way to the dispute. Based on this and other information, the requesting party substantially revises its list of requested key words and phrases to eliminate certain terms that appear to generate "junk" results. Further sampling of the revised query results, which are much smaller than before, suggests that more than half of the records retrieved are likely relevant to the dispute.

In situations involving very large databases or multiple databases, test queries or pilot tests of the production process can be based on a subset of the data repository, consistent with the approach outlined in *Zubulake v. UBS Warburg LLC*⁵⁷ and elsewhere. Although the *Zubulake* opinions do not concern database information, the court's approach of using small, manageable test queries to generate empirical results from which the burden and benefit of further discovery could be determined has been widely adopted in other ESI situations, including discovery of database information.

Sharing technical or logistical information and using sampling to more effectively negotiate the scope of discovery are also consistent with guidance contained in *The Sedona Conference Commentary on Achieving Quality in the E-Discovery Process* (2013) and *The Sedona Conference Cooperation Proclamation* (2008).⁵⁸

4. Validation

A responding party should use reasonable measures to validate that its collection from the database is both reasonably complete and did not inadvertently modify the ESI.

Comment 4.

Due to the volume of information and the complexities of its organization inside databases, there are no established protocols or integrity checks (e.g., MD5 hash marking) to verify and validate the completeness and accuracy of database information collected from a larger database. However, verifying that information extracted from databases is an accurate copy of the same information as it is stored in the original database should not be seen as an insurmountable task; as a matter of due diligence, basic checks exist to ensure the completeness, accuracy, and integrity of the collected data.

Extracting data from a database in response to a discovery request typically involves: (1) executing a query to identify responsive records; and (2) structuring the responsive fields into an export format acceptable for production. Running queries and structuring output files frequently can result in unintended changes to data values, such as truncating text, substituting codes for values, or other data transformations. Other typical data extraction problems include unintentionally extracting records that are not responsive (over-inclusion) or missing records that should be included (under-inclusion) in the production set. These and other data integrity issues can render the resulting dataset incomplete or inaccurate, and thus unacceptable for production.

To reduce the risk that information extracted from databases contains transcription errors, a responding party that is extracting data from a database and formatting it into a report or file for the purpose of responding to a discovery request should test the proposed dataset to confirm that it includes all expected content and complies with the target format.

⁵⁷ *Zubulake*, 217 F.R.D. 309, *supra*.

⁵⁸ However, a responding party is not obligated to run test queries and provide sampling information to requesting parties to satisfy curiosity. For example, when a responding party reasonably believes that a database or other structured data source contains no relevant information, it should not be obligated to sample the system absent particularized and credible evidence to the contrary. See Principle 6: Responsibilities of Responding Parties, *supra*, at II.C. See *The Sedona Conference Commentary on Achieving Quality in E-Discovery*, 15 Sedona Conf. J. 265 (2014) and see *The Sedona Conference Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009 Supp.).

Depending on the nature of the data and the methodology used to extract the data, a variety of validation procedures may be considered:

- **Validating numeric values.** When data consists of a numeric value, the following tests may be appropriate:
 - Confirm that the number of extracted database records matches the number of records that were originally identified by one or more target queries.
 - Compare the resulting number of records to the number that appears in reports that are regularly produced in the ordinary course of business.
 - Compare the number of extracted records to control counts from the tables being queried.
 - Compare the aggregate of certain fields, such as sales amounts, to known control totals from routine or regularly produced reports.
 - Develop control totals by confirming that the sum total of the extracted records plus the total of the non-extracted records equals the total of the same field or record set as noted in the entire table or report.

Illustration vi. A party requests information about all buyers of a product, including the date of purchase, the price paid, and the state in which the purchase took place. All of this information is tracked in a sales database maintained by the responding party. The responding party runs a query to identify all the sales records for that specific product and exports the requested information into a .CSV file. Before producing this information, the responding party double-checks the number of data rows in the .CSV file by loading it into a spreadsheet program and comparing the number of lines to the number of records identified in the query. The responding party then checks to make sure that the date and price field rows contain only date and price information. Satisfied with the results of these checks, the responding party then provides this information to the requesting party.

- **Validating standard language values.** When the contents of an extracted field consist of standard language values rather than a numeric value, the responding party should confirm that the extracted text values conform to a list of expected values for those fields. For example, for fields that can contain only a limited number of valid values, such as the seven days of the week or the twelve months of the year, a responding party can run an automated comparison of the extracted information against all possible expected values for these fields to ensure that no unexpected values are included.
- **Validating non-standardized language values.** When text fields do not require standardized language, as in many narrative or comment fields, a sample of fields from the extracted text can be examined to confirm that

the information meets expectations of the information that should be stored there. Samples of extracted text fields can also be compared to the corresponding records in routine or regularly prepared reports to confirm that the extracted text field information is consistent with presentation of the same information in validated reports used in the ordinary course of business.

- **Validating from multiple fields.** In situations where values in the production dataset are calculated from several fields in the source database, responding parties can help make the extracted fields more easily validated by including not only the field containing a calculated result field value, but also the source field values from which the resultant values are calculated. Including this additional information would make it possible for both requesting and responding parties to check the internal consistency of the final result field.
- **Validating from multiple tables (relational databases).** In relational databases, multiple tables of data are often linked by key values that are echoed on one or more tables. Extracted database information that has either been retrieved from or is being produced in multiple tables can be checked for accuracy and completeness by confirming that the linking key values from the various tables are consistent and sufficient to properly link the records from the various tables. Ambiguous key values – i.e., values that do not provide a unique relationship between correct data elements – can occur when information is extracted from multiple tables.
- **Validating from reports.** Finally, responding parties should not underestimate the ability of database reports in general to confirm the accuracy of a data extraction. Many standard reports that are used on a regular basis within an organization, including regulatory filings generated through queries or scripted tools, compile sophisticated information and metrics that can be used to double-check the accuracy and consistency of many types of data fields extracted from a database.

Authenticating exported database information builds on validation processes, and more than one procedure can be used to demonstrate sufficient consistency, completeness, and accuracy in the extracted data. However, situations can occur in which field values are different in the source database and in the extracted data. Typically, such differences are caused by mechanical issues, such as a report template that truncates the information in a field after the first N characters, thereby displaying only a partial entry that cannot be fully validated against the original database input. However, if these differences are not caught soon after the extracted data has been prepared and produced, the consequences of relying upon the extracted data can have far-reaching consequences. Both requesting and producing parties should consider adding quality assurance procedures to ensure that such errors are quickly identified.

5. Data Authenticity and Admissibility

The proper validation of collection from a database does not automatically make the substantive information stored in the database authentic, admissible or true. These are separate issues that need to be analyzed by the appropriate decision makers.

Comment 5.A. Causes of Inaccuracy in Database Information

Although businesses may rely on database data or reports in the ordinary course of business, the fact that data is derived from a database does not make it any more intrinsically reliable than other types of evidence produced in discovery. Databases, whether simple or complex, are not infallible. The “true” accuracy of the underlying data depends on many factors. Systems or components can malfunction, errors may occur in programs and formulas, manual data entry may introduce errors, and certain cells, fields or tables can be mislabeled or misinterpreted (e.g., a table of numbers reflecting a certain volume of widgets sold could pertain to either individual widgets or units of widgets, if values are not properly labeled or represented by a credible witness with knowledge). In addition, as mentioned previously, the way that certain fields within a database are used may change over time, meaning that old data records and new data records may use the same fields but record different information. Sometimes, current users of the database are not even aware of these changes.

While rare, it is also possible that a responding party or its counsel may have intentionally or unintentionally manipulated database output in a way that degrades the quality of the data being produced. Such degradation may take the form of data that lacks certain metadata fields that are integral to understanding the remainder of the information.⁵⁹

Comment 5.B. Standards for Admitting Database Information into Evidence

Because the production of information extracted from databases may be composed of different elements – e.g., raw data, individual data cells, printed summary reports – the lack of consistency can make the process of authenticating the substantive content of this information a complex task. While there are currently no bright-line rules for authentication of database information, several opinions suggest that tests for admissibility of database information are becoming more stringent.⁶⁰ Discussion and application of the Federal Rules of Evidence are beyond the scope of this Commentary; however, *The Sedona Conference Commentary on ESI Evidence & Admissibility*⁶¹ offers useful analyses of cases that reflect the various “evidentiary hurdles” that a proponent seeking to admit electronically stored information into evidence must clear.⁶²

Across and even within jurisdictions, there is significant disparity between the most lenient and most demanding approaches for admitting database information as substantive evidence. While some of this disparity also may take into account proportionality considerations, parties seeking to make use of database information should be prepared to establish a rigorous foundation for this evidence. For the party that produced the database information, this may require calling one or more witnesses to trial

59 See *Bray & Gillespie Mgmt. LLC v. Lexington Ins. Co.*, 259 F.R.D. 568, 585-87 (M.D. Fla. 2009) (reviewing case law for proposition that production of static TIFF images of email from database, stripping all metadata fields, violated Rule 34, *vacated in part*, 2009 WL 5606058 (M.D. Fla. Nov. 16, 2009) (reversing Magistrate Judge’s finding that attorneys had acted in bad faith).

60 See, e.g., *In re Vee Vinhnee*, 2005 WL 3609376, 06 Cal. Daily Op. Serv. 146, 2006 Daily Journal D.A.R. 169 (B.A.P. 9th Cir. Dec 16, 2005) (detailing factors that impact admissibility of database information); *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534 (D. Md. 2007).

61 *The Sedona Conference Commentary on ESI Evidence & Admissibility*, 9 Sedona Conf. J. 217 (2008).

62 One 9th Circuit opinion, *U-Haul Int’l, Inc. v. Lumbermens Mutual Casualty Co.*, 576 F.3d 1040 (9th Cir. Aug. 12, 2009), affirmed a more lenient standard in analyzing the district court’s admissibility of computer-generated summaries of payments made on insurance claims. Finding that such summaries were properly admitted, the appellate court focused primarily on the four basic steps of the business records exception to hearsay under Fed. R. Evid. 803(6): (1) the underlying data was entered into the database at or near the time of each payment event; (2) the persons who entered the data had knowledge of the payment event; (3) the data was kept in the course of Republic Western’s regularly conducted business activity; and (4) the claims manager was qualified and testified as to this information. 576 F.3d. at 1044-45.

who can establish the foundation. For other parties, this may require deposing a representative of the producing party. While it may not require every single *Vinbnee* factor,⁶³ an evidentiary proffer of database information may require a witness who can explain the origins and lifecycle of the information in the ordinary course of business, as well as the procedures used to extract this data and prepare an exhibit of this information for trial.⁶⁴ Litigation-specific exhibits, as opposed to copies of reports or database views used in the ordinary course of business, are likely to draw special attention from both opponents and the presiding court, as validation procedures used to double-check business reports may not have been applied to litigation-driven work. Parties should consider reducing cost and saving trial time by stipulating to admissibility, where appropriate.

6. Form of Production

The way in which a requesting party intends to use database information is an important factor in determining an appropriate format of production.

Comment 6.A. Discussing the Intended Reasonable and Legitimate Uses of Database Information Can Result in a More Useful Production Format

While a requesting party is not required to divulge its counsel's work product or its litigation strategy, it may be impossible for a responding party to take appropriate steps to provide database information in a reasonably useful format if it has no idea of how the requesting party intends to use it. A requesting party's failure or refusal to identify the intended use of database information, especially upon request, may limit the responding party's ability to accommodate the format request, particularly where the responding party's preferred format is less expensive and appears *ex ante* reasonable. To maximize the value of the database information it will receive, a requesting party should provide detail sufficient to describe the tools or broad evidentiary use that it intends to make of this material. For example, a party's desire to review some database information in conjunction with witnesses' statements or testimony may make a database report the most useful way of receiving this information. Other times, a requesting party may wish to analyze or otherwise manipulate the database information to show relationships within the data. Disclosing the specific database or analytical engine that a requesting party intends to use – without revealing the precise type of analysis that will take place – enables the responding party to make reasonable efforts to accommodate the requestor's proportional, reasonable, and legitimate uses of the data, and thus better understand the technical specifications required for the production. To the extent that the parties cannot resolve questions of appropriate production format, this level of information also will facilitate a swift and appropriate decision by the court.

Like relevance, any assessment of a requesting party's stated "reasonable and legitimate use" of database information should provide sufficient latitude so that requesting parties can conduct their litigation as they generally see fit. However, the mere fact that databases contain large amounts of information does not permit a party to submit broad

⁶³ *In re Vee Vinbnee, supra.*

⁶⁴ *Santander Consumer USA, Inc. v. Superior Pontiac Buick GMC, Inc.* No. 10–13181, 2012 WL 5363553 (E. D. Mich., Oct. 30, 2012) (rejecting argument that database records are inadmissible hearsay where a party does not own the database and finding that the plaintiff presented a witness who was familiar enough with the database system and record-keeping process to satisfy Rule 803(6)). Contrast with *Meyer Corporation U.S. v. Alfay Designs*, No. CV 20103647(CBA)(MDG), 2012 WL 3536987 (E.D. N.Y. Aug. 13, 2013) (imposing sanctions on a party who, despite knowing the depth of technical knowledge required for a deponent, produced an employee who could not even answer basic questions about the database system and retention policies).

discovery requests merely to satisfy idle curiosity or to use data beyond what is necessary to prosecute alleged claims and defenses.

Comment 6.B. *Factors for Determining Reasonableness of Data Production Format*

Under Rule 34(b)(2)(E)(ii), if a requesting party does not specify the form of production, the data must be produced “in a form or forms in which it is ordinarily maintained or in a reasonably usable form or forms.” And Rule 34(a) contemplates, “translation by the responding party into a reasonably usable form,” “if necessary.” The Committee Notes for the 2006 Amendments to Rule 34(b) explain that whether a responding party is required to convert information to a “more usable form, or should be required to produce it at all, should be addressed under Rule 26(b)(2)(B) [proportionality factors].” The Notes also make it clear that responding parties are not allowed to produce the information in a form “that makes it more difficult or burdensome for the requesting party to use the information efficiently in the litigation.”

Thus, Under Rule 34 and the accompanying 2006 Advisory Committee Note, the key factors for reasonableness of production format include whether there is any loss of information from the original format and whether the requesting party can make appropriate use of the database information. The Note also points to a third factor – proportionality (as measured under Rule 26 and its state analogs) – that also should be part of the analysis. A request for database information that requires a disproportionate amount of effort from the responding party should not be permitted, even if a lesser response does not provide the same degree of information access as would have the initial request.⁶⁵

Illustration vii. Requesting party seeks all of the invoice records of Company X’s billing system from 2002-2006. The requesting party plans to use them as exhibits at trial, but it wants to easily search and find the specific invoices. Because the requesting party will not be using this database information to perform trend or other relational analysis, a searchable production of the invoices as fixed image files may be reasonably usable, provided sufficient searching information for the invoices is provided.

Illustration viii. Requesting party seeks all of the invoice records of Company X’s billing system from 2002-2006. The requesting party plans not only to use individual invoices as exhibits at trial, but it also wants to analyze aggregated invoice information by customer over time to see whether the Company has a pattern of double billing after the fourth invoice. Because the requesting party intends to use the data to undertake legitimate relational analysis, a fixed-imagined production may not be reasonably useable.

Producing database information in a reasonably usable form neither requires a responding party to produce it in a format that is the best or optimum format for the requesting party, nor ensures that such data requires little or no manipulation by a responding party. If the effort, ability, and cost to transform the data into a specific requested format are similar for both the requesting and responding parties, a strong argument can be made that the requesting party should bear the cost, so long as the initial production format was, in and of itself, reasonable.

⁶⁵ See *The Sedona Conference Commentary on Proportionality in Electronic Discovery*, 14 SEDONA CONF. J. 155 (2013).

Illustration ix. In a small-dollar contract dispute, the requesting party asks for invoice data stored in a database to be produced in table format with each row constituting a single invoice to a single customer. The responding party does not have direct access to the database and cannot easily run custom data extractions from the database. Instead, the responding party's built-in reporting script can create individualized invoices that contain identical data, but not in tabular form. The best technology available to the parties involves scanning invoices to manually create tables of information. The producing party argues that the requesting party should bear the cost of further manipulation of the data, as the production of searchable individualized invoices was reasonable, given the amount in controversy, the lack of information lost by the production format, and the equal burden for both requesting and responding parties.

THE SEDONA CONFERENCE BEST PRACTICES COMMENTARY ON THE USE OF SEARCH & INFORMATION RETRIEVAL METHODS IN E-DISCOVERY*

A Project of The Sedona Conference Working Group on Electronic Document Retention & Production (WG1)

Author:

The Sedona Conference

2013 Editors-in-Chief:

Jason R. Baron

Maura R. Grossman

2007 Editor-in-Chief:

Jason R. Baron

2007 Senior Editors:

Thomas Y. Allman

M. James Daley

George L. Paul

Thanks go to all who participated in the dialogue that led to this Commentary.

We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

PREFACE AND ACKNOWLEDGEMENTS

Welcome to the 2013 Edition of The Sedona Conference *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*. Since the publication of the 2007 Public Comment Version of this Search Commentary, there have been significant developments in both case law and technology in the area of search and retrieval. Indeed, the 2007 Search Commentary itself has been prominently cited in a number of reported cases as an authoritative source on best practices in this area.

The 2013 Edition of the Commentary reflects changes in legal practice with a new section on computer- or technology-assisted review, as well as citations to more recent case law. Certain of the original eight Practice Points have been revised to reflect developments in law and practice, including recognition of the key principles of cooperation and proportionality advanced by The Sedona Conference. The Appendix on Information Retrieval Methods has also been modified to reflect changes in technology. The text of the 2007 Version of this Commentary otherwise remains largely intact, except for the deletion and/or updating of outdated information, and for minor stylistic and grammatical edits. The text was not edited with an eye towards being a fully-revised “Second Edition” of the original Commentary. Nevertheless, The Sedona Conference recognizes that the rapidly-evolving nature of automated techniques calls for continuing close attention to further changes in professional practice in this area, especially with respect to defending the process used, and we will endeavor to meet that need through future publications.

I want to thank the entire Working Group for all their hard work and contributions, and especially the 2013 Editorial Committee for leading this effort to update the existing Search Commentary. I wish to acknowledge the contributions of Jason R. Baron and Maura R. Grossman for taking the lead in revising and updating the prior version, as assisted by Bobbi Basille, Todd Elmer, Amir Milo, Priya Keshav, and James Sherer. Finally, but certainly not least, the Working Groups of The Sedona Conference could not accomplish their goals without the financial support of the sustaining and annual sponsors of the Working Group Series listed at www.thesedonaconference.org/sponsors.

Kenneth J. Withers
Deputy Executive Director
The Sedona Conference
December 2013

TABLE OF CONTENTS

Overview.....220

Executive Summary.....222

 I. Introduction.....226

 II. The Search and Information Retrieval Problem Confronting Lawyers228

 III. Lawyer’s Current use of Search and Retrieval Methodologies230

 IV. Some Key Terms, Concepts, and History in Information Retrieval Technology237

 V. Boolean and Beyond: A World of Search Methods, Tools, and Techniques241

 VI. Practical Guidance for Evaluating the use of Automated Search and Retrieval Methods243

 VII. Future Directions in Search and Retrieval Science249

Appendix: Types of Search Methods255

OVERVIEW

***Traditional Approaches To Searching For Relevant Evidence
Are No Longer Practical Or Financially Feasible***

Discovery of relevant information about a topic in dispute is at the core of the litigation process.¹ However, the advent of “e-discovery” is causing a rapid transformation in how that information is gathered. While discovery disputes are not new, the huge volume of available electronically stored information (“ESI”) poses unique challenges. Some years ago, a party facing a review of information for production to the other side in a document-intensive case might have been concerned with hundreds of “banker’s” boxes of documents.

Today, that same amount of data is easily found on a single computer hard drive.² Moreover, as the ability to create and store massive volumes of electronic information mushrooms, the cost to store that information inversely plummets. In 1990, a gigabyte of storage cost about \$20,000; as of 2013, two-terabyte drives readily sell for less than \$70, or 3.5¢ per gigabyte, with even lesser rates charged for hosting gigabytes in the “cloud.” As a result, more individuals and organizations are generating, receiving, and storing more data, which in some cases means more information must be identified, collected, reviewed, and produced in litigation.

With billable rates for associates at many law firms averaging between \$200 and \$500 per hour,³ the cost to review just one gigabyte of data can easily exceed \$30,000.⁴ These economic realities – i.e., the huge cost differential between the nominal cost to store a gigabyte of data and the \$30,000 to review it – act as drivers to change traditional attitudes and approaches of lawyers, clients, courts, and litigation support providers forced to search for relevant evidence during discovery and investigations. Data volumes now numbering in the billions of ESI objects, review costs, and shrinking discovery timetables have created the need for a profound change in practice.

As discussed in this Commentary, just as technology has given rise to these new litigation challenges, technology can help to solve them. The emergence of new discovery strategies, best practices, and processes, as well as new search and retrieval technologies are transforming the way lawyers litigate. Collectively, they provide opportunities for huge volumes of information to be reviewed faster, more accurately, and more affordably than ever before. The good news is that search and retrieval systems are improving in effectiveness and expanding their capacities, buoyed by a tsunami of activity aimed at improving the “search” experience for users generally.⁵ For example, advanced forms of

1 *Hickman v. Taylor*, 329 U.S. 495, 507 (1947) (“Mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.”).

2 Here’s why: One gigabyte of electronic information can generate approximately 70,000 to 80,000 pages of text, or 35 to 40 banker’s boxes of documents (at 2,000 pages per box). Thus, a 250 gigabyte storage device (e.g., a laptop or hard drive), theoretically, could hold as much as the equivalent of 8,750 to 10,000 banker’s boxes of documents. In contrast, in 1990, a typical personal computer held just 200 megabytes of data – less than 1/1000 the capacity of a typical hard drive today. Even if only 10% of a computer’s available capacity today contains user-created information (as distinguished from application programs, operating systems, utilities, etc.), attorneys still would need to consider and potentially review 1,750,000 to 2,000,000 pages per device.

3 See Alex Vorro, *Law firm billing rates steadily climbing despite down economy*, (April 17, 2012) (citing TyMetrix Legal Analytics’ 2012 Real Rate Report”), <http://www.insidecounsel.com/2012/04/17/law-firm-billing-rates-steadily-climbing-despite-d>.

4 See *infra* note 18 and accompanying text.

5 One indication of the amount of ongoing effort and investment to improve search and retrieval capabilities is evidenced by the research and development spending of internet and technology giants Google, Microsoft, Apple, and IBM. According to published reports, Google spent \$ 3.76 billion, Microsoft spent \$8.7 billion, Apple spent \$1.78 billion, and IBM spent \$6.03 billion on core research and development activities in 2010. See Booz & Company, *The 2011 Global Innovation 1000 – Why Culture Is Key* (October, 2011), <http://www.booz.com/media/file/BoozCo-Global-Innovation-1000-2011-Webinar.pdf>.

machine learning – including supervised and unsupervised document and content classifiers – can automatically organize ESI in new ways not achieved by the more familiar methods of the past (which include the use of simple “keywords” as an automated aid to conducting manual searches). And not only can these new techniques increase accuracy and efficiency, through the proper use of statistical sampling and validation techniques, practitioners can measure the accuracy of the results of either traditional or alternative forms of search, retrieval, and review.

New challenges require new solutions. This Commentary aspires to present the bench and bar with an intelligible picture of the new challenges associated with the search and retrieval of ESI. The Commentary also presents alternative ways to address those challenges and to select the best solution for a given set of circumstances, taking into account the just, speedy, and inexpensive determination of every action (consistent with Federal Rule of Civil Procedure 1).

EXECUTIVE SUMMARY

Discovery has changed. For a growing number of cases, the process of identifying, reviewing, and producing information has evolved from the manual review of paper documents to an evaluation of vastly greater volumes of ESI.

A perfect review of the resulting volume of information is impossible. It is also not economical. But governing legal principles and best practices do not require perfection in making disclosures or in responding to discovery requests.⁶ Instead, best practices focus on reasonable and proportional actions taken by practitioners as part of their duties, which must include an appreciation for the particular challenges of electronic information.

The Sedona Conference has helped establish the benchmarks governing the evolution and refinement of reasonable, good faith practices for searching intimidating amounts of data. Principle 6 of *The Sedona Principles, Second Edition* (2007) notes that “[r]esponding parties are best situated to evaluate the procedures, methodologies and technologies appropriate for preserving and producing their own electronically stored information,” and Principle 11 amplifies the point by stating that “[a] responding party may satisfy its good faith obligation to preserve and produce relevant electronically stored information by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information.”

This Commentary discusses the existing and evolving methods by which a party may choose to search unprecedented volumes of information. As the practice of using these “search and retrieval” technologies – the generic term we will utilize in this Commentary – advances, a new understanding of what is “reasonable” and “proportional” under any particular set of circumstances will advance as well. Therefore, the challenges addressed by this Commentary go beyond litigation and also encompass the full breadth of the search and retrieval of information from large volumes of data.

The Revolution in Discovery

Not long ago, all information was stored on physical records such as paper. There was typically a single “original” document, and the number of duplicate copies and their locations were generally limited. Administrative assistants, file clerks, records managers, and archivists developed expertise in managing that storage, generally pursuant to pre-existing file systems. In the case of litigation, it was reasonable and relatively easy (in all but the exceptional case), for a legal representative to gather, manually review, and prepare each individual item prior to its production.

The digital revolution did more than make documents truly portable – it also created a review-process paradigm shift in terms of what is truly feasible regarding document review in litigation. This revolution has shifted nearly all information storage (as a percentage of existing information) to the digital realm and has caused an explosion in the amount of information that resides in any organization. And not only did the information’s volume and format change, the very geography of where information “lives” moved from a file cabinet to a broad distribution amongst many different storage devices: from large mainframe computers to handheld devices.

⁶ See, e.g., *Pension Comm. of Univ. of Montreal Pension Plan v. Banc of Am. Sec., LLC*, 685 F. Supp. 2d 456, 461, 465 (S.D.N.Y. 2010), overruled in part on other grounds by *Chin v. Port Authority of New York & New Jersey*, 685 F.3d 135 (2nd Cir. 2012).

Each device may be capable of storing the equivalent of several warehouses of paper documents; each device may also have networking capabilities which allow it to integrate into complex systems. These systems are intricate, interdependent, and evolve spontaneously, behaving nearly like living ecosystems. To further complicate the picture, a legal professional who completely understands the workings of this new form of “information ecosystem” is rare indeed.

Finally, in addition to the search and retrieval challenge, a large percentage of the records searched in litigation are written in human language, not just numbers. Human language is an inherently elastic, ambiguous, “living” tool of enormous power. Its elasticity allows for jargon, private codes, and discrete vocabularies to exist in different subcultures in any organization, thereby making the identification of search terms much more challenging.

Essential Conclusions of this Commentary

This Sedona Conference *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery* strives to set forth state-of-the-art knowledge defining the challenges associated with searching enormous databases for relevant information (the “Problems”), and presents methods and tools to retrieve that information with a minimum of wasted effort (the “Solutions”).

By way of summary, we set forth our conclusions about the Problems and their Solutions, and summarize our Practical Advice articulated in the balance of this Commentary.

Problems

- The exponential growth in digital information is a critical challenge to the justice system.
- Parties are frequently unable to identify ESI that is likely to contain information relevant to the claims and defenses in the dispute.
- Electronically stored information consists of human language, which challenges computer search tools. These challenges are posed by the ambiguity inherent in human language; the imprecision resident in human use of logic; and the tendency of people within organizations or networks to speak in metaphor, to invent their own words, and to communicate in jargon, short forms, or code.
- The application of simple keyword searching, while still a valuable tool, has well-documented deficiencies. There are also documented problems with manual document review.

Solutions

- Educating clients that good information governance reduces e-discovery costs by reducing the volume of ESI that is kept, and effective management of the ESI that is kept results in collecting a smaller collection with a higher concentration of relevant information.

- Counsel should work closely with their clients to identify and then narrow sources of ESI that are likely to contain information relevant to the dispute.
- The proper selection of information for production in discovery can benefit from the learning from a variety of other disciplines, including, but not limited to, Information Retrieval science, linguistics, and the implementation of effective project management processes.
- Alternative search tools may properly supplement simple keyword searching and Boolean search techniques. These include using various forms of computer- or technology-assisted review, machine learning, relevance ranking, and text mining tools which employ mathematical probabilities, as well as other techniques incorporating supervised and unsupervised document and content classifiers.⁷
- Parties and their counsel should cooperate and seek ways to agree on measurements to evaluate the effectiveness of the search and retrieval process. The metrics currently used in information retrieval science, most notably “precision” and “recall,” may serve as key points of reference.

Practical Advice

- Practice Point 1.* *In many settings involving large amounts of relevant electronically stored information (“ESI”), relying solely on a manual search process for the purpose of finding responsive documents may be infeasible or unwarranted. In such cases, the use of automated search methods should be viewed as reasonable, valuable, and even necessary under certain circumstances.*
- Practice Point 2.* *The successful use of any automated search method or technology will be enhanced by a well-thought-out process with substantial human input on the front end.*
- Practice Point 3.* *The choice of a specific search and retrieval method will be highly dependent on the specific legal context in which it is to be employed. Parties and their counsel must match the use case with the tools and best practices appropriate to address it, and must incorporate proportionality considerations involving the overall costs and the stakes of the litigation.*
- Practice Point 4.* *Parties and their counsel should perform due diligence when choosing a particular information retrieval product or vendor service.*

⁷ There is great variation in the description and application of these technologies, whether for technical, sales, or marketing differentiation or other business purposes. Some of the terms currently in use as of the date of this Commentary include: computer-assisted review; technology-assisted review; predictive coding; relevance ranking; text mining; tools that employ mathematical probabilities; as well as other techniques, including fuzzy logic to capture variations on words, and conceptual search, which makes use of taxonomies and ontologies assembled by linguistic means. For a glossary of terms relating to these technologies see Maura R. Grossman and Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review with Foreword by John M. Facciola*, U.S. Magistrate Judge, 2013, FED. CTS. L. REV. 7 (January 2013), <http://www.fclr.org/fclr/articles/html/2010/grossman.pdf>.

- Practice Point 5. Because of the characteristics of human language, no search and information retrieval tool can guarantee the identification of all responsive documents in large data collections. Moreover, different search methods may produce different results, subject to a measure of statistical variation inherent in the science of information retrieval.*
- Practice Point 6. Parties and their counsel should make a good faith attempt to cooperate when determining the use of particular search and information retrieval methods, tools, and protocols (including keywords, concepts, computer- or technology-assisted review and other types of search parameters and quality control measures.*
- Practice Point 7. Parties and their counsel should expect that their choice of search methodology (and any validation of it) will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and at trial).*
- Practice Point 8. Parties, counsel, and the courts should be alert to new and rapidly evolving search and information retrieval methods. Moreover, parties and their counsel should recognize that information retrieval is a distinct field of study that includes expertise in such areas as computer science, statistics, and linguistics, and that consultation with or utilization of experts in information retrieval may improve the quality of search results in complex cases involving large volumes of ESI.*

How The Legal Community Can Contribute to The Growth of Knowledge

A consensus is forming in the legal community that human review of documents in discovery is expensive, time consuming, and error-prone. There is also a growing awareness that, used correctly, linguistic and mathematically-based content analysis, embodied in new forms of search and retrieval technologies, tools, techniques, and processes in support of the review function, can effectively reduce litigation cost, time, and error rates.

Recommendations

- 1. The legal community should continue to support collaborative research with the scientific and academic sectors aimed at establishing the efficacy of a range of automated search and information retrieval methods.*
- 2. The legal community should encourage the establishment of objective benchmarking criteria, to assist lawyers in their evaluation of the competitive legal and regulatory search and retrieval services market.*

Members of The Sedona Conference community have and will continue to participate in collaborative workshops and other forums dedicated to information retrieval issues. The Sedona Conference intends to remain in the forefront of the efforts of the legal community aimed at seeking out centers of excellence in this area, including the possibility of fostering private-public partnerships focusing on continued research.

I. INTRODUCTION

The exponential growth in the volume and complexity of ESI found in modern organizations poses a substantial challenge to the justice system. Today, even routine discovery requests can require searches of, and retrieval from, the storage devices found on servers, networked workstations, desktops and laptops, home computers, removable media (such as CDs, DVDs, and USB flash drives), handheld devices (such as PDAs, smart phones, cell phones, and iPods), and the “cloud.” Complicating things further, such information is now almost always flowing robustly throughout a “network,” in which it has likely been replicated, distributed, modified, linked, attached, accessed, backed-up, overwritten, deleted, undeleted, fragmented, defragmented, morphed, and multiplied. Complying with preservation or discovery obligations in some ESI cases may require a process to identify relevant emails from among thousands, millions, or even tens-of-millions of individual messages, with attachments in various file formats.

The volume and complexity of ESI highlights several issues: First, whether automated search and information retrieval methods are reliable and accurate, and if so, how accurate. Second, whether the legal profession has the skills, knowledge, and processes required to use such automated search and retrieval methods intelligently in conjunction with huge data sets, in ways that are defensible under the rules governing discovery.

In *The Sedona Principles*, Second Edition (2007), The Sedona Conference endorsed several highly pragmatic and relevant consensus best practices relevant to this discussion.⁸

First, Principle 6 provides that parties responding to discovery are in the best position “to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.” Principle 11 expands this concept to include the use of “electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information.”

Second, the Commentary to Principle 11 provides that the “selective use of keyword searches can be a reasonable approach when dealing with large amounts of electronic data,” and states that it “is also possible to use technology to search for ‘concepts,’ which can be based on ontologies, taxonomies, or data clustering approaches, for example.”⁹ This exploits a unique feature of electronic information: the ability to conduct fast, iterative searches for the presence of patterns of words and concepts in large document populations. The Commentary to Principle 11 also states that “[c]ourts should encourage and promote the use of search and retrieval techniques in appropriate circumstances,” and suggests that “[i]deally, the parties should agree on the search methods, including search terms or concepts, to be used as early as practicable. Such agreements should take account of the iterative nature of the discovery process and allow for refinement as the parties’ understanding of the relevant issues develops.”¹⁰

Third, The Sedona Conference has recognized that “there are now hundreds of companies offering electronic discovery services.”¹¹ This is also true of search and

⁸ The Sedona Principles, *Best Practices Recommendations & Principles for Addressing Electronic Document Production* (2d ed. 2007) (hereinafter *The Sedona Principles*), available at <https://thesedonaconference.org/download-pub/81>.

⁹ *Id.* at Comment 11.a.

¹⁰ *Id.*

¹¹ The Sedona Conference Best Practices for the Selection of Electronic Discovery Vendors: *Navigating the Vendor Proposal Process* (2007), available at <https://thesedonaconference.org/download-pub/80>.

information retrieval products and services for use in legal contexts – which form a subset of a burgeoning sector of the economy devoted to improving users’ “search” experience. However, there remains some confusion as to the strengths and weaknesses of such tools. Legal practitioners have a need for guidance as to the appropriate use of search and information retrieval technologies. Such guidance can help practitioners judge the relative costs and benefits of such tools in specific cases.

This Commentary is designed to help educate the justice system – attorneys, judges, and litigants alike – on “state of the art” search and retrieval tools, techniques, and methodologies, and how they can best be used as part of an overall process to more efficiently manage discovery. This discussion includes the critically important concept of an integrated process of search and retrieval; the ability to differentiate among different search methods; how to evaluate such differences; and what questions to ask before using any particular method or product in a specific legal setting.

For the past three decades, the legal community has had familiarity with simple keyword and natural language searches on Westlaw[®] and Lexis[®] in the context of legal research, and to a lesser extent the use of “Boolean” logic to combine keywords and “operators” (such as “AND,” “OR,” and “AND NOT,” or “BUT NOT”) that produce broader or narrower searches. Over time, lawyers have applied this knowledge to employ simple keyword, Boolean, and other search and retrieval tools to reduce the amount of information to be reviewed for production in discovery.¹² In the past few years, the relative efficacy of competing search and retrieval tools used to accomplish review for production has begun to be measured. However, the field is still wide open for the development of more advanced search and information retrieval best practices. These methods include merging keyword searching with more sophisticated systems that use computer- and technology-assisted techniques, incorporating mathematical algorithms and various forms of linguistic techniques, to help find, group, and present related content.

What follows is an in-depth analysis of the problems lawyers confront in managing massive amounts of data in discovery, including how search and retrieval techniques are used in everyday practice and the key element of “process.” This Commentary also provides background on the field of information retrieval and at least partially describes the world of search tools, techniques, and methodologies that are currently commercially available. It also includes “practice pointers” on the factors to consider in making an overall legal evaluation among different search methods, both on a conceptual and practical level. In a concluding section, the future of search and retrieval efforts is discussed. A more technical discussion of various search methodologies is included in an Appendix. Where appropriate, reference will be made to technical definitions found in The Sedona Conference Glossary: *E-Discovery and Digital Information Management* (4th ed. 2014).¹³

12 There may be a role for the use of some type of search and retrieval technology in discharging obligations to preserve ESI, as well as during the initial pre-review data culling or “collection” phases, in anticipation of complying with specific ESI and document requests. During the collection phase, for example, the goal is to maximize the amount of potentially relevant evidence in a subset of the greater universe of available ESI, without necessarily selecting only the more relevant information that might be the focus of the review phase preceding production. Accordingly, parties may well end up using (and agreeing to use) differing search methods in the initial collection and later review phases of litigation. While we acknowledge that use of advanced search tools during earlier phases of litigation (e.g., during early case assessment, at preservation, etc.) remains cutting edge and worthy of future discussion, the primary focus of this Commentary will be on search tools as they are used in the review process. See generally Thomas Y. Allman, Jason R. Baron, and Maura R. Grossman, *Preservation, Search Technology, and Rulemaking*, 30 THE COMPUTER AND INTERNET LAWYER No. 2 (February 2013); Mia Mazza, Emmalena K. Quesada, and Ashley L. Sternberg, *In Pursuit of FRCP 1: Creative Approaches To Cutting and Shifting the Costs of Discovery of Electronically Stored Information*, 13 RICH. J. L. & TECH. 11 (2007), at [53] & [60], available at <http://law.richmond.edu/jolt/v13i3/article11.pdf> (discussing the use of concept searching in regard to preservation); *The Sedona Principles*, supra n.8 (“Organizations should internally address search terms and other filtering criteria as soon as possible so that they can begin a dialogue on search methods as early as the initial discovery conference.”).

13 The Sedona Conference Glossary: *E-Discovery & Digital Information Management* (4th ed. 2014), 15 SEDONA CONF. J. 305 (2014).

II. THE SEARCH AND INFORMATION RETRIEVAL PROBLEM CONFRONTING LAWYERS

Discovery today is drowning in an exponential flood of potential sources of information. This increase in volume, especially since the mid-1990s, is principally due to the combined effects of the PC revolution, the widespread use of email and other new forms of communication, and the growth of mobile device and social networks. Indeed, the implication of this growth in volume is that it places at severe risk the justice system's ability to achieve the "just, speedy and inexpensive" resolution of disputes, as contemplated by Rule 1 of the Federal Rules of Civil Procedure.

The Rise of Crushing Volumes of Information in the Digital Realm

A history of the computer and information technology advances occurring since the mid-1970s is beyond the scope of this Commentary. Suffice it to say that over the last 40 years, there has been a fast-paced and widespread shift from physical information storage technologies to new, digital information storage technologies. This "digital realm" was created by an accretion of technological advances, each built on preceding advances, which together have resulted in as fundamental a shift in the way information is shared, such as that which occurred in 1450 when Johannes Guttenberg invented the printing press. Included among the advances contributing to the new "digital realm" are the invention of the microchip, the development and diffusion of the personal computer, the spread of various types of networks linking together both computers and other networks, the rise of email and its dominant use in the business world, the plunging cost of computing power and storage, and of course, the spread of the Internet and with it, the World Wide Web.¹⁴

By the mid-1990s, networked computers and their storage devices had created a true information-based society, with a constant flow of messages in all forms exchanged on a 24/7 basis. For example, studies reflect that the typical corporate worker sends and receives about 105 emails per day.¹⁵ The size and nature of the attachments to these emails is also growing, with increased integration of image, audio, and video files. More recently, there has been a similar explosion in the use of instant and text messaging throughout organizations, including increasingly, through the use of mobile devices. In many organizations, the average worker maintains several gigabytes of stored data.¹⁶ At the same time, the costs of storage have plummeted from \$20,000 per gigabyte in 1990 to less than 3.5¢ per gigabyte in 2013.¹⁷ Existing technologies are only beginning to grapple with providing a viable automated means for applying records retention requirements, including the ability to implement legal holds, in the new digital world.

14 See George L. Paul and Jason R. Baron, *Information Inflation: Can the Legal System Adapt?*, 13 RICH. J. L. & TECH. 10 (2007), at [1], n.2, available at <http://law.richmond.edu/jolt/v13i3/article10.pdf> ("Organizations have thousands if not tens of thousands of times as much information within their boundaries as they did 20 years ago."); Peter Lyman and Hal R. Varian, *How Much Information*, 2003, available at <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>.

15 See Sara Radicati, Ed., *Email Statistics Report, 2011-2015*, The Radicati Group, Inc., May 2011, available at <http://www.radicati.com/wp/wp-content/uploads/2011/05/Email-Statistics-Report-2011-2015-Executive-Summary.pdf>.

16 One gigabyte is equivalent in volume to between 70,000 and 80,000 pages of material. At 2,000 pages per box, one gigabyte is therefore equivalent to 35 to 40 boxes of documents. See *supra* n.2.

17 See *Memory Storage Density*, Wikipedia, http://en.wikipedia.org/wiki/Memory_storage_density#Effects_on_price (last visited April 7, 2013); Michelle Kessler, *Days of officially drowning in data almost upon us*, USA TODAY, Mar. 5, 2007, www.usatoday.com/tech/news/20070305data_N.htm. Cloud storage reduces the costs of storage even further. See, e.g., <https://developers.google.com/storage/docs/pricingandterms> (last visited April 7, 2013); <http://mashable.com/2012/08/21/amazon-glacier> (last visited April 7, 2013).

Organizations have continued to aggressively leverage technology to increase productivity. But leveraged technology sometimes comes with a lack of oversight or control. In many organizations, no one really controls how, where, how many times, and in how many forms information is stored. For example, copies of the same Word document can be found in email attachments, local hard drives, network drives, document management systems, websites, and on all manner of backup and removable media, such as USB flash drives, CDs, DVDs, and so on.

Discovery In the Recent Past: Manageable Amounts of Physically Stored Information

Historically, outside counsel played a key role in a comparatively simpler discovery process: Litigants, assisted by their counsel, identified and collected information that was relevant to pending or reasonably foreseeable litigation. Counsel manually reviewed the information and produced any information that was responsive and not otherwise protected from disclosure by the attorney-client privilege, the attorney work product, or by trade secret protections.

This worked fine in the days where most of the potentially relevant information was created in or was stored in printed, physical form, and in reasonable volumes, so that it required only “human eyes” to review and interpret it. However, with increasingly complex computer networks, and the exponential increase in the volume of information existing in the digital realm, the venerated process of “eyes only” review is no longer generally workable or economically feasible.

The cost of manual review of such volumes is prohibitive, often exceeding the damages at stake. Anecdotal reports indicate that the cost of reviewing information can easily exceed thousands of dollars per custodian, per event, for collection and attorney review. Litigants often cannot afford to review all available ESI in the time permitted for discovery.¹⁸ Accordingly, the conventional document review process is poorly adapted to a growing percentage of today’s litigation.¹⁹ Lawyers of all stripes therefore have a vital interest in utilizing automated search and retrieval tools where appropriate. The plaintiff’s bar has a particular interest in being able to efficiently extract key information received in mammoth document productions, and in automated tools that facilitate the process. The defense bar has an obvious interest in reducing attendant costs, increasing efficiency, and in better risk management of litigation (including reducing surprises). All lawyers, clients, and judges have an interest in reducing cost and barriers to entry to the justice system, and maximizing the quality of discovery, by means of using automated tools that produce a reliable, reproducible, and consistent result.

Ideally, then, judges and litigants should strive to increase their awareness of search and retrieval sciences generally, and of the sciences’ appropriate application to discovery. Some technologies have been used for years to produce documents from large litigant document databases, but often without much critical analysis. The legal system may benefit from the rich body of research available through the Information Retrieval and library science disciplines. The discussion that follows is designed to provide a common framework

18 Compare mere pennies to store a gigabyte of data with \$32,000 to review it in a traditional, linear fashion (i.e., assuming one gigabyte equals 80,000 pages and assuming that an associate billing \$200 per hour can review 50 documents per hour at 10 pages in length, such a review would take 160 hours at \$200/hr., or approximately \$32,000). See generally Nicholas M. Pace & Laura Zakaras, WHERE THE MONEY GOES: UNDERSTANDING LITIGANT EXPENDITURES FOR PRODUCING ELECTRONIC DISCOVERY 17-27 (2012) [hereinafter RAND 2012 STUDY], <http://www.rand.org/pubs/monographs/MG1208.html>.

19 Not all cases are equally reliant on electronic discovery – from time to time, counsel may even forgo the production of electronically stored information and rely solely on hard-copy documents.

and vocabulary for proper application of search and retrieval technologies in this new “age of information complexity” in the legal environment.

The Reigning Myth of “Perfect” Retrieval Using Traditional Means

It is not possible to discuss this issue without noting that there appears to be a myth that manual review by humans of large amounts of information is as accurate and complete as possible – perhaps even perfect – and constitutes the gold standard by which all searches should be measured. Even assuming that the profession had the time and resources to continue to conduct manual or “linear” review of massive sets of electronic data (which it does not), the relative efficacy of that approach versus utilizing newly developed automated methods of review appears to be increasingly in question.²⁰ Moreover, past research demonstrates the gap between lawyers’ expectations and the true efficacy of certain types of keyword searches. The Blair and Maron study (discussed below) shows that human beings are far less accurate and complete than they believe themselves to be when searching and retrieving information from a heterogeneous set of documents (i.e., in many data types and formats), using ad hoc, simple keywords as the sole means to identify potentially relevant documents. The importance of this point cannot be overstated, as it provides a critical frame of reference in evaluating how new and enhanced forms of automated search methods and tools may benefit litigation practices.

The Intelligent Use of Tools

Although the continued use of manual search and review methods may be infeasible or even indefensible in discovery involving significant amounts of ESI, merely adopting sophisticated automated search tools, alone, will not necessarily lead to successful results. Lawyers must recognize that the process by which a legal team uses such tools, including close involvement of lead counsel, is just as important as the automated tools themselves. This may require an iterative process which importantly incorporates feedback and learning and allows for measurement and validation of results. The time and effort spent up front on designing a sophisticated discovery process that targets the real needs of the litigation must be viewed as a condition precedent to deploying automated methods of search and retrieval.

III. LAWYERS’ CURRENT USE OF SEARCH AND RETRIEVAL METHODOLOGIES

Attorneys across all disciplines are generally familiar with search and retrieval methodologies based on their exposure over the past thirty-plus years to automated means of searching of caselaw and other databases provided by LexisNexis’ and Westlaw’. More recently, lawyers are using Google’ and other Web-based search engines to hunt down the increasing amounts of online information relevant to their practices. Additionally, law firms and corporate legal departments use search methods for administrative matters, such as locating data on personnel, supporting billing functions, managing conflicts of interest, and for contact management. Many products employing search methods of various kinds exist in the legal marketplace to assist lawyers in these functions.

Current Database Tools in the Practice of Law

Litigators use automated search and retrieval tools at many stages of the litigation process. PACER and other automated means are used to uncover data on opposing

²⁰ See *infra* text accompanying note 48.

counsel's pleadings, motions, and pretrial filings in similar litigation, as well as showing how a judge has ruled on similar issues even if unreported in legal reporting services. Lawyers also use a variety of search methods involving online, CD-ROM, client-developed, and "cloud" databases to unearth facts on opposing parties, witnesses, and even potential jurors. At later stages of litigation, lawyers use various litigation support software applications to search through potential exhibits in connection with proceedings held in "electronic courtrooms." But until recently, litigators seldom used automated search and retrieval methods with their clients' or their opponents' growing collections of unstructured ESI.

"Deduplication" in the Processing of ESI

With the exponential increase in the volume of data subject to e-discovery, lawyers have begun to take steps towards employing automated search tools to manage the discovery process. One example of this is "deduplication" software used to find duplicate electronic files, since ESI often consists of a massively redundant universe. For example, the same email can be copied tens or even hundreds of times in different file locations on a network or on backup media. Deduplication software reduces the time attorneys must spend reviewing a large document set and helps to ensure consistent classification of documents for responsiveness or privilege.²¹ Increasingly, "email threading" and "near deduplication" tools are used to assist in organizing and expediting overall document reviews, even if the technique is not used to reduce the actual number of unique documents subject to review.²²

The Use of "Keywords"

The most commonly used search methodology today still entails the use of "keyword searches" of full text and metadata as a means of filtering data for producing responsive documents in civil discovery. For the purpose of this Commentary, the use of the term "keyword searches" refers to set-based searching using simple words or word combinations, with or without Boolean and related operators (see below and Appendix for definitions). The ability to perform keyword searches against large quantities of evidence has become a widely accepted practice, as recognized by the courts. As one United States Magistrate Judge stated in 2004, "the glory of electronic information is not merely that it saves space but that it permits the computer to search for words or 'strings' of text in seconds."²³

-
- 21 "Deduplication" tools tag identical documents as duplicates by means of a "binary hash function" (a mathematical way of comparing the text of two documents represented in the underlying digital 1's and 0's actually stored on the computer to see if the documents are perfectly alike). Deduplication by binary hash has been widely used without much notice in court opinions to date. See *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 571 (N.D. Ill. 2004) (referring to deduplication process); *Medtronic Sofamor Danek Inc. v. Michelson*, 229 F.R.D. 550, 561 (W.D. Tenn. 2003) (same).
- 22 "Email threading" refers to a particular message and a running list of all subsequent replies pertaining to that original email. David D. Lewis & Kimberly A. Knowles, *Threading electronic mail: A preliminary study* 33 INFORMATION PROCESSING AND MANAGEMENT 209 – 217 (1997) ("Near deduplication" involves files that "are not hash value duplicates but are materially similar."); available at http://pdf.aminer.org/000/936/211/threading_electronic_mail_a_preliminary_study.pdf.
- 23 *In re Lorazepam & Clorazepate*, 300 F. Supp. 2d 43, 46 (D.D.C. 2004); see also *In re CV Therapeutics, Inc.*, 2006 WL 2458720 (N.D. Cal. Aug. 22, 2006) (court endorses employment of search terms as reasonable means of narrowing production); *J.C. Associates v. Fidelity & Guaranty Ins. Co.*, 2006 WL 1445173 (D.D.C. 2006) (requiring search of files using four specified keywords); *FTC v. Ameridebt, Inc.*, 2006 WL 618563 (N.D. Cal. Mar. 13, 2006) ("email could likely be screened efficiently through the use of electronic search terms that the parties agree upon"); *Windy City Innovations, LLC v. America Online, Inc.*, 2006 WL 2224057 (N.D. Ill. July 31, 2006) ("[k]eyword searching permits a party to search a document for a specific word more efficiently"); *Reino de Espana v. Am. Bureau of Shipping*, 2006 WL 3208579 (S.D.N.Y. Nov. 3, 2006) (court approves of keyword search for names and email addresses as a "targeted and focused discovery search"); *U.S. ex rel. Tyson v. Amerigroup Ill., Inc.*, 2005 WL 3111972 (N.D. Ill. Oct. 21, 2005) (referencing agreement by parties to search terms); *Medtronic*, 229 F.R.D. at 561 (court orders defendant to conduct searches using the keyword search terms provided by plaintiff); *Alexander v. FBI*, 194 F.R.D. 316 (D.D.C. 2000) (court places limitations on the scope of plaintiffs' proposed keywords to be used to search White House email).

Courts have not only accepted, but in many cases have ordered, the use of keyword searches to define discovery parameters and resolve discovery disputes.²⁴ Early on, one court suggested that a party might satisfy its duty to preserve documents in anticipation of litigation by conducting a system-wide keyword search and preserving a copy of each “hit.”²⁵

Because of the costs and burdens associated with the review of increasingly vast volumes of electronic data, it makes sense in appropriate cases for producing parties to negotiate with requesting parties in advance to define the scope of discoverable information. For example, parties could agree on conducting a search of only files maintained by key witnesses, in certain data sources, and/or for certain date ranges. They may negotiate and agree to a set of key words relevant to the case. Both sides might see the advantage to using such protocols or filters to reduce the volume of extraneous information, such as spam, routine listserv notifications, and personal correspondence, typically found when searching through electronic data collections.²⁶

In *Treppel v. Biovail Corp.*,²⁷ the defendant refused to produce documents because the plaintiff would not agree to keyword search terms. Citing to Principle 11 of the Sedona Principles for Electronic Document Production, the Court held that the defendant was justified in using keyword search terms to find responsive documents and should have proceeded unilaterally to use its list of terms when the plaintiff refused to endorse the list. The Court held that plaintiff’s “recalcitrance” did not excuse defendant’s failure to produce any records and ordered the company immediately to conduct the automated search, produce the results, and explain its search protocol. Another early case emphasized the need to confer after plaintiff was successful in obtaining a “mirror image” of data on all of defendant’s computers.²⁸

Issues With Keywords

There are nonetheless a number of notable limitations to the effectiveness of traditional or basic keyword searching. Keyword searches work best when the legal inquiry is focused on finding particular documents and when the use of language is relatively predictable. For example, basic keyword searches work well to find all documents that mention a specific individual or date, regardless of context. However, while basic keyword searching techniques have been widely accepted both by courts and parties as sufficient to

24 See generally Jason R. Baron, *Law in the Age of Exabytes: Some Further Thoughts on ‘Information Inflation’ and Current Issues in E-Discovery Search*, 17 RICH. J.L. & TECH. 9 (2011) (compiling case law), available at <http://jolt.richmond.edu/v17i3/article9.pdf>.

25 *Zubulake v. UBS Warburg, LLC*, 229 F.R.D. 422 (S.D.N.Y. 2004); see also *Zakre v. Norddeutsche Landesbank Girozentrale*, 2004 WL 764895 (S.D.N.Y. Apr. 9, 2004) (court denies plaintiff’s request for additional indexing of records, holding that defendant’s production of CD-ROMS in a text searchable form was sufficient, citing to Guideline 11 of *The Sedona Principles*, 2004 Edition). cf. *Cache La Poudre Feeds, LLC v. Land O’Lakes, Inc.*, 2007 WL 684001 (D. Colo. Mar. 2, 2007) (where court denied motion for sanctions based on an allegation that the opposing party failed to properly monitor compliance with its discovery obligations by not conducting keyword searches, court also stated that *The Sedona Principles*, 2004 Edition and *Zubulake* were not to the contrary).

26 See also R. Brownstone, *Collaborative Navigation of the Stormy e-Discovery Seas*, 10 RICH. J.L. & TECH. 53 (2004), available at <http://law.richmond.edu/jolt/v10i5/article53.pdf> (arguing that parties must agree to search terms and other selection criteria to narrow the scope to manageable data sets); *The Sedona Principles*, supra note 8, Comment 11.1.a (“For example, use of search terms could reveal that a very low percentage of files (such as emails and attachments) on a data tape contain terms that are responsive to ‘key’ terms. This may weigh heavily against a need to further search that source, or it may be a factor in a cost-shifting analysis. Such techniques may also reveal substantial redundancy between sources (i.e., duplicate data is found in both locations) such that it is reasonable for the organization to preserve and produce data from only one of the sources.”). See generally Kenneth J. Withers, *Computer-Based Discovery in Federal Court Litigation*, 2000 FED. CTS. L. REV. 2 (2000), <http://www.fcltr.org/fcltr/articles/html/2000/fedctslrev2.shtml> (suggesting parties adopt collaborative strategies on search protocols).

27 233 F.R.D. 363 (S.D.N.Y. 2006).

28 *Balboa Threadworks v. Stucky*, 2006 WL 763668 (D. Kan. Mar. 24, 2006) (court ordered parties to meet and confer on the use of a search protocol, including keyword searching).

define the scope of their obligation to perform a search for responsive documents, the experience of many litigators is that simple keyword searching alone is inadequate in at least some discovery contexts. This is because simple keyword searches are both over- and under-inclusive in light of the inherent malleability and ambiguity of spoken and written English (as well as all other languages).²⁹

Traditional keyword searches identify all documents containing a specified term regardless of context, often capturing many documents irrelevant to the user's query. For example, the term "strike" could be found in documents relating to a labor union tactic, a military action, options trading, or baseball, to name just a few (illustrating "polysemy," or ambiguity in the use of language). The problem of the relative percentage of "false positives" or noise in the data is potentially huge, amounting in some cases to enormous numbers of files which must be searched to find responsive documents.³⁰

On the other hand, basic keyword searches have the potential to miss documents that contain a word that has the same meaning as the term used in the query, but is not specified. For example, a user making queries about labor actions might miss an email referring to a "boycott" if that particular word was not included as a keyword, and a lawyer investigating tax fraud via options trading might miss an email referring to "exercise price" if that term was not specifically searched (illustrating "synonymy" or variation in the use of language). And of course, if authors of records are inventing words "on the fly," or using short-forms or code names (as they have done throughout history, and with increasing frequency in electronic communications), such problems are compounded.³¹

Keyword searches can also exclude common or inadvertently misspelled instances of the term (e.g., "Phillip" for "Philip," or "strik" for "strike") or variations on "stems" of words (e.g., "striking"). Even the best of optical character recognition (OCR) programs introduce a certain rate of random error into document texts, potentially transforming would-be keywords into gibberish. Finally, using keywords alone results in a return set of potentially responsive documents that are not weighted or ranked based upon their potential importance or relevance. In other words, each document is considered to have an equal probability of being responsive subject to further manual review.

More advanced keyword searches using "Boolean" operators and techniques borrowed from "fuzzy logic" may increase the number of relevant documents and decrease the number of irrelevant documents retrieved. These searches attempt to emulate the way humans use language to describe concepts. In essence, they simply translate ordinary words and phrases into a Boolean search argument. Thus, a natural language search for "all birds that live in Africa" is translated to something like ("bird* + liv* + Africa").

At the present time, it would appear that the majority of automated litigation support providers and software rely on some form of keyword searching, although the legal landscape is changing (see discussion below). Such methods are limited by their dependence

29 Some case law has held that keyword searches were either incomplete or over inclusive. See *Alexander v. FBI*, 194 F.R.D. at 316; *Quinby v. WestLB AG*, 2006 WL 2597900 (S.D.N.Y. Sept. 5, 2006) (court narrows party's demand for 170 proposed search terms in part due to the inclusion of commonly used words).

30 See, e.g., G. Paul and J. Baron, *Information Inflation*, *supra* n.14, at [20] (discussing potential time and cost of searching through 1 billion emails); Craig Ball, *Crafting A More Effective Keyword Search*, LAW TECHNOLOGY NEWS (June 24, 2009), <http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202431693400&slreturn=20130302113909>.

31 Philosophers use colorful imagery to describe the dynamism and complexity of human language. See, e.g., Ludwig Wittgenstein, *THE PHILOSOPHICAL INVESTIGATIONS*, Section 18 (G.E.M. Anscombe, trans., 3d ed. 1973) ("[T]o imagine a language is to imagine a form of life. ... [L]anguage can be seen as an ancient city; a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this surrounded by a multitude of new boroughs with straight regular streets and uniform houses").

on matching a specific, sometimes arbitrary choice of language to describe the targeted topic of interest.³²

However, these challenges can be at least partially overcome by employing a more methodical and informed approach to defining keywords. Such a process begins with a clear definition of relevance, outlining criteria to identify relevant documents for each issue and subtopic. The problem of false positives can be minimized by combining key terms within a certain proximity of one another or in a specified order. Singular keywords are often ambiguous, but disambiguating (for verbs) and specifying words (for nouns) when joined to the central keyword with Boolean operators can reduce over-inclusiveness. Additionally, gaps in keywords can be a big issue in early stages or when the issues at stake are relatively unknown. One approach to identify such gaps trains a software system on initial custodians (i.e., utilizes a software-assisted review workflow) and uses the trained system to generate a set of potential keywords that can then be used to confirm or add to the original assumptions regarding keywords in subsequent stages of review.³³

Recent judicial opinions, including several citing to the 2007 Version of this Commentary, have examined many of the limitations of keyword searching.³⁴ Notably, the Court in *William A. Gross Construction Associates* declared a

[W]ake-up call to the Bar ... about the need for careful thought, quality control, testing, and cooperation with opposing counsel in designing search terms or 'keywords' to be used to produce e-mails or other electronically stored information ('ESI'). ... Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords, with input from the ESI's custodians as to the words and abbreviations they use, and the proposed methodology must be quality control tested to assure accuracy in retrieval and elimination of 'false positives.'³⁵

Use of Alternative Search Tools and Methods

Lawyers are beginning to feel more comfortable using alternative search tools to identify potentially relevant ESI. These more advanced text mining tools include, but are not limited to, "conceptual search methods," which rely on semantic relations between words, and/or use "thesauri" to capture documents that would be missed in keyword searching. Specific types of alternate search methods are set out in detail in the Appendix.

"Concept" search and similar information retrieval technologies attempt to locate information that relates to a desired concept without the presence of a particular word or phrase. The classic example is the concept search that will recognize that documents about Eskimos and igloos are related to Alaska, even if they do not specifically mention the word

32 See *infra* Part IV; see generally S.I. Hayakawa, LANGUAGE IN THOUGHT AND ACTION (5th ed.1990) (stating that such methods are inherently limited by their specific choice of language to describe a specific object or reality).

33 See generally The Sedona Conference *Commentary on Achieving Quality in E-Discovery* (2013), 15 Sedona Conf. J. 265 (2014), (discussing how to construct a better quality search process).

34 See, e.g., *United States v. O'Keefe*, 537 F. Supp. 2d 14, 24 (D.D.C. 2008) (Facciola, M.J.); *Equity Analytics, LLC v. Lundin*, 248 F.R.D. 331, 333 (D.D.C. 2008) (Facciola, M.J.); *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 260, 262 (D. Md. 2008) (Grimm, M.J.); see also *NDLON v. Immigration and Customs Enforcement, Dept of Homeland Security*, 811F Supp. 2d 713 (S.D.N.Y. 2011) (Scheidlin, J.); see generally Baron, *Law in the Age of Exabytes*, *supra* note 24.

35 *William A. Gross Construction Associates, Inc. v. American Manufacturers Mutual Insurance Co.*, 256 F.R.D.265, 134, 136 (S.D.N.Y. 2009) (Peck, M.J.); see also Hon. Andrew Peck, *Search, Forward*, LAW TECH. NEWS (Oct. 1, 2011) at 1-2, available at http://www.recommind.com/sites/default/files/LTN_Search_Forward_Peck_Recommind.pdf.

“Alaska.” The first reported case referencing the possible use of “concept searching” as an alternative to strict reliance on keyword searching was decided in 2007.³⁶

Other automated tools rely on “taxonomies” and “ontologies” to help find documents conceptually related to the topic being searched, based on commercially available data or on specifically compiled information. This information is provided by attorneys or developed for the business function or specific industry (e.g., the concept of “strike” in labor law vs. “strike” in options trading). These tools rely on the information that linguists collect from the lawyers and witnesses about the key factual issues in the case – the people, organization, and key concepts relating to the business as well as the idiosyncratic forms of communication that might be lurking in documents, files, and emails. For example, a linguist would want to know how union organizers or company officials might communicate plans, any special code words or lingo used in the industry, the relationships of collective bargaining units, the company’s management structure, and other issues and concepts.

Another type of search tool relies on mathematical probabilities that a certain text is associated with a particular conceptual category. These types of machine learning tools, which include “clustering” and “latent semantic indexing,” are potentially helpful in addressing cultural biases of taxonomies because they do not depend on linguistic analysis, but on mathematical probabilities. They can also help identify communications hidden in code language and neologisms. For example, if the labor lawyer were searching for evidence that management was targeting neophytes in the union, she might miss the term “n00b” (a neologism for “newbie”). This technology, first used in government intelligence, and now increasingly used as part of computer- or technology-assisted review in e-discovery, is particularly apt in helping lawyers find information when they do not know exactly what to look for. For example, when a lawyer is looking for evidence that key players conspired to violate the labor union laws, she will usually not know the “code words” or expressions the players may have used to disguise their communications. For a discussion of recent developments in computer- or technology-assisted review, see below.

With so many different search methods currently available, it is important to choose the most appropriate search strategy for any particular case. The choice of a search method will always depend heavily on the particular context. Practitioners should be aware of the strengths and limitations of varying approaches. Sometimes the most appropriate search method is obvious from the outset of a case; in other situations, the best method(s) only become evident after experimentation and use. But practitioners must recognize if a particular search method is ineffective and must be willing to modify their approach based on the results. See also Practice Pointers, below.

Resistance by the Legal Profession

Some litigators continue to primarily rely upon manual review of information as part of their review process.³⁷ Principal rationales are: (1) concerns that computers cannot be trusted to replace the human intelligence required to make complex determinations on

³⁶ See *Disability Rights Council of Greater Washington v. Washington Metro. Transit Auth.*, 242 F.R.D. 139, 148 (D.D.C. 2007); see generally *Mazza, et al.*, *supra* note 12, at [54] (discussing concept searching).

³⁷ *But see In re Insinet Group, Inc.*, 2005 WL 3501708 at *3 (Del. Ch. Dec. 14, 2005). The court reduced plaintiffs’ attorneys’ fee claim by \$1 million (75% of the total claim) for “obvious” inefficiencies in plaintiffs’ counsel’s review of paper printouts (“blowbacks”) from digital files. The court stated that plaintiffs’ counsel’s decision to “blow back” the digital documents to paper “both added unnecessary expense and greatly increased the number of hours required to search and review the document production.”

relevance and privilege; (2) the perception that there is a lack of scientific validity of search technologies necessary to defend against a court challenge; and (3) widespread lack of knowledge (and confusion) about the capabilities of automated search tools.

Other parties and litigators may accept simple keyword search, yet be reluctant to use alternative search techniques. They may not be convinced that the chosen method would be defensible if confronted with a court challenge. They may perceive a risk that problem documents will not be found despite the additional effort, or that documents might be missed which would otherwise be picked up in a straight keyword search. Moreover, acknowledging that there is no one solution for all situations, they may opt for an accepted, lowest common denominator approach. Finally, litigators often lack the time and resources to sort out these highly complex technical issues on a case-by-case basis.³⁸

But the legal landscape is changing rapidly. The year 2012 saw the first judicial opinions approving the use of the alternative search method of computer- or technology-assisted review (as described in Section V, below).³⁹ In the *Moore* opinion, the magistrate judge noted that “[c]ounsel no longer have to worry about being the “first” or “guinea pig” for judicial acceptance of computer-assisted review. ... Computer-assisted review now can be considered judicially-approved for use in appropriate cases.”⁴⁰ This and future precedent will hasten acceptance by lawyers (and their clients) of the use of such alternate methods and techniques.

Challenging the Choice of Search Method

Challenge to the choice of a search methodology used in a review prior to production can arise in one of two contexts: (1) a requesting party’s objection to the unilateral selection of a search method by a responding party; or (2) a court’s *sua sponte* review of the use of a method or technology. Accordingly, the preferable method to preempt challenges – advocated by the proponents of the 2006 Federal Rules Amendments and some practitioners – is for a full and transparent discussion among counsel of the search methodology. Where the parties are in agreement on the method and a reasonable explanation can be provided, it is unlikely that a court will second-guess the process.

Absent agreement, a party has the presumption, under Sedona Principle 6, that it is in the best position to choose an appropriate method of searching and culling its data. However, a unilateral choice of a search methodology may risk challenge if an opponent can show that the results of the search are not accurate, complete, or reliable. As a practical matter, those who might object to a particular search and retrieval technology may face several challenges. First, the legal system has, for decades, blessed the use of keyword search tools and databases for discovery review. And second, if human review or even keyword searching is the benchmark for accuracy and completeness, it arguably should not be

38 See, e.g., Ron Friedmann, *A Future Beyond Hammers*, STRATEGIC LEGAL TECHNOLOGY BLOG (Feb. 4, 2005, 1:38PM), <http://prismlegal.com/a-future-beyond-hammers/> (suggesting that not one solution fits all cases); see also Ron Friedmann, *Thoughts on Full Text Retrieval (A KM and Litigation Support Topic)*, STRATEGIC LEGAL TECHNOLOGY BLOG (July 30, 2003) (questioning the incremental value of sophisticated searching over simple searching because of the costs of implementation and need to build taxonomies and to test methodologies).

39 See *Moore v. Publicis Groupe SA*, 287 F.R.D. 182 (S.D.N.Y. 2012) (Peck, M.J.), *aff’d*, *Moore v. Publicis Groupe SA*, 2012 WL 1446534 (S.D.N.Y. Apr. 26, 2012) (Carter, J.) (approving joint search protocol for technology-assisted review); see also *In re Actos (Pioglitazone) Products*, 2012 WL 3899669 (W.D. La. July 27, 2012) (same); *Global Aerospace Inc. v. Landow Aviation LP*, Consol. Case No. CL 00061040-00 (Va. Cir. Ct. Apr. 23, 2012) (Chamblin, J.) (granting responding party’s request to use technology-assisted review); *EORHB, Inc. v HOA Holdings, LLC*, Civ. No. 7409-VCL (Del. Ch. Oct. 15, 2012) (ordering, *sua sponte*, parties to use technology-assisted review with same vendor).

40 *Moore*, 287 F.R.D. at 195.

difficult to measure new technologies against keyword searching or human review, especially when guided by a reasonable process. The discovery standard is, after all, reasonableness, not perfection.

Given the continued exponential growth in information, a large body of precedent will likely develop over time that critically analyzes new and alternative search methods in use in particular legal contexts. Indeed, the first case in which a court held an evidentiary hearing on a challenge to the use of keywords in favor of one type of alternative computer-assisted review method occurred in 2012.⁴¹

IV. SOME KEY TERMS, CONCEPTS, AND HISTORY IN INFORMATION RETRIEVAL TECHNOLOGY

The evaluation of Information Retrieval (“IR”) systems has, at least until recently, largely been of interest to computer scientists and graduate students in information and library science. Unlike performance benchmarking for computer hardware, there are no accepted, objective criteria for evaluating the performance of IR systems. That is, for IR systems, the notion of effectiveness is subjective. Human judgment is ultimately the criterion for evaluating whether an IR system returns the relevant information in the correct manner. Two users may have differing needs when using an IR system. For example, one may want to find all potentially relevant documents. Another may want to correctly sort information by priority. In addition, subject matter and information type may impact a user’s IR requirements.

Over the past 50 years, a large body of research has emerged concerning the evaluation of IR systems. The study of IR metrics helps quantify and compare the benefits of various search and IR systems. In 1966, C.W. Cleverdon listed various “metrics” which have become the standard for evaluating IR systems within what has become known as the “Cranfield tradition.”⁴² Two of the metrics, *precision* and *recall*, are based on binary relationships. That is, either a document is relevant or it is not, and either a document is retrieved or it is not. Several modifications and additional metrics have been added in the IR literature since then, as the scientific field continues to add and refine techniques for measuring the efficiency of IR systems – both in terms of retrieval and also in user access to relevant information.

Measuring the Effectiveness of Information Retrieval Methods

Recall, by definition, is “an information retrieval performance measure that quantifies the fraction of known relevant documents which were effectively retrieved.”⁴³ That is, out of the total number of relevant documents in the document collection, how many were retrieved correctly?

41 See *Kleen Prod., LLC v. Packaging Corp. of Am.*, 2012 WL 4498465 (N.D. Ill. Sept. 28, 2012) (Nolan, M.J.) (discussing evidentiary hearings held prior to parties reaching agreement on choice of method).

42 See Cyril W. Cleverdon, et al., ASLIB CRANFIELD RESEARCH PROJECT: FACTORS DETERMINING THE PERFORMANCE OF INDEXING SYSTEMS (Vol. I., 1966), available at [http://www.sigir.org/museum/pdfs/Factors_Determining_the_Performance_of_Indexing_Systems_Vol_2/pdfs/frontmatter.pdf](http://www.sigir.org/museum/pdfs/Factors%20Determining%20the%20Performance%20of%20Indexing%20Systems%20Vol%20I%20-%20Part%20I%20Text/pdfs/frontmatter.pdf); Cyril W. Cleverdon, et al., ASLIB CRANFIELD RESEARCH PROJECT: REPORT OF CRANFIELD II (Vol II, 1966), available at http://www.sigir.org/museum/pdfs/Factors_Determining_the_Performance_of_Indexing_Systems_Vol_2/pdfs/frontmatter.pdf; see generally, C.J. Rjiisbergen, INFORMATION RETRIEVAL (2d ed. 1979), available at <http://www.dcs.gla.ac.uk/Keith/Preface.html> (last visited November 25, 2013).

Precision is defined as “an information retrieval performance measure that quantifies the fraction of retrieved documents which are known to be relevant.”⁴⁴ Put another way, how much of the returned result set is “on target”?

Recall and precision can be expressed by simple ratios:

$$\text{Recall} = \frac{\text{Number of responsive documents retrieved}}{\text{Number of responsive documents overall}}$$

$$\text{Precision} = \frac{\text{Number of responsive documents retrieved}}{\text{Number of documents retrieved}}$$

If a collection of documents contains, for example, 1,000 documents, 100 of which are relevant to a particular topic and 900 of which are not, then a system that returned only these 100 documents in response to a query would have a precision of 1.0, and a recall of 1.0.

If the system returned all 100 of these documents, but also returned 50 of the irrelevant documents, then it would have a precision $100/150 = .667$, and still have a recall of $100/100 = 1.0$.

If it returned only 90 of the relevant documents along with 50 irrelevant documents, then it would have a precision of $90/140 = 0.64$, and a recall of $90/100 = 0.9$.

Importantly for the practitioner, there is typically a trade-off between precision and recall. One can often adjust a system to retrieve more documents – increasing recall – but the system achieves this result at the expense of retrieving more irrelevant documents – decreasing precision. Effectively, one can cast either a narrow net and retrieve fewer relevant documents, along with fewer irrelevant documents, or cast a broader net and retrieve more relevant documents, but at the expense of retrieving more irrelevant documents.⁴⁵

Measuring the Efficiency of Information Retrieval Methods

Efficiency is important to the success of an IR system, but it does not affect the quality of the results. Efficiency is measured in three ways. The first measurement is the mean time for returning search results (this can be measured by the average time it takes to return results, or the computational complexity of the search). The second measurement is the mean time it takes a user to complete a search. This measurement is more subjective and is a function of the ease of use of the IR system. A third method involves the number of documents that must be reviewed to achieve a particular level of recall or precision.

⁴³ See Ricardo Baeza Yates & Berthier Ribeiro Neto, MODERN INFORMATION RETRIEVAL 437, 455 (1999) (glossary), available at <http://www.sims.berkeley.edu/~hearsst/irbook/glossary.html>; see also *The Grossman-Cormack Glossary*, *supra* n.7.

⁴⁴ R. Yates & B. Neto, *supra* n.43, at 455.

⁴⁵ There are many other common metrics that are considered in information retrieval literature, including F-measure, average precision, and average search length. F-measure is an approximation of the “crossover point” between precision and recall, where both are maximized. Average precision determines the precision level for each retrieved relevant item. Average search length is the average position of a relevant retrieved item. Still other terms include “fallout” (the ratio of the number of non-

The Blair and Maron Study

A well-known study testing recall and precision in a legal setting was conducted by David Blair and M.E. Maron in 1985.⁴⁶ The Blair and Maron study demonstrated the problems caused by the rich use of human language among the many people that can be involved in a dispute and how difficult it is to take such richness into account in a search for information. Indeed, Blair and Maron found that attorneys were only about 20% effective at identifying all of the different ways that document authors could refer to words, ideas, or issues in their case.

For the purposes of their study, Blair and Maron evaluated a case involving an accident on the San Francisco Bay Area Rapid Transit (BART) in which a computerized BART train failed to stop at the end of the line. There were about 40,000 documents, totaling about 350,000 pages, in the discovery database. The attorneys worked with experienced paralegal search specialists in an effort to find all of the documents that were relevant to the issues. The attorneys estimated that they had found more than 75% of the relevant documents, but more detailed analysis found that the number was actually only about 20%. The authors found that the different parties in the case used different words to describe the same thing, depending on their role in the case. The parties on the BART side of the case referred to “the unfortunate incident,” but parties on the victim’s side called it a “disaster.” Other documents referred to the “event,” “incident,” “situation,” “problem,” or “difficulty.” Proper names were often not mentioned.

As Roitblat notes, *supra*, note 46, Blair and Maron even found:

that the terms used to discuss one of the potentially faulty parts varied greatly depending on where in the country the document was written. Some people called it an ‘air truck,’ a ‘trap correction,’ ‘wire warp,’ or ‘Roman circle method.’ After 40 hours of following a ‘trail of linguistic creativity’ and finding many more examples, Blair and Maron gave up trying to identify all of the different ways in which the document authors had identified this particular item. They did not run out of alternatives, they only ran out of time.

More Recent Studies on Precision and Recall in E-Discovery

In the years since publication of the 2007 Version of this Commentary, a variety of other efforts have been made to study the precision/recall issues in a legal discovery context. Many of these have been initiated by members of The Sedona Conference. Numerous studies emanating from the TREC Legal Track (see discussion, below) have confirmed relatively low rates of recall obtained from basic keyword searching.⁴⁷ Moreover, two widely-cited recent studies have provided the foundation for lawyers making the claim that some of the more advanced computer- or technology-assisted review methods yield more accurate results than reliance solely on human review (as measured by precision,

relevant items retrieved to the total number of items retrieved”) and “elusion” (the proportion of responsive documents that have been missed by the search). See generally, the Grossman-Cormack Glossary, *supra* n.7.

⁴⁶ David L. Blair & M.E. Maron, *An Evaluation of Retrieval Effectiveness For a Full-Text Document-Retrieval System*, 28 Comm. ACM 289 (1985). The discussion that follows of the Blair and Maron study is drawn directly from Herbert L. Roitblat, *Search and Information Retrieval Science*, 8 SEDONA CONF. J. 225, 231 (2007).

⁴⁷ See, e.g., Stephen Tomlinson, et al., *Overview of The TREC 2007 Legal Track*, http://trec.nist.gov/pubs/trec16/t16_proceedings.html (baseline Boolean search captured only 22% of universe of all relevant

recall, and F_1 measures).⁴⁸ Both the 2010 study by Roitblat et al., and the 2011 study by Grossman and Cormack, used secondary data to compare the effectiveness of human review to that of certain computer- or technology-assisted review methods. Roitblat et al. used a collection of 1.6 million documents reviewed in response to a Department of Justice Second Request. The 1.6 million documents were re-reviewed by two teams using (unspecified) technology-assisted review methods. A statistical sample of 5,000 documents from the 1.6 million was also re-reviewed by two teams using manual review. The two technology-assisted review efforts both yielded better agreement with the original production than the two human review efforts, supporting the conclusion that technology-assisted review can be at least as effective as human review. Grossman and Cormack used the TREC 2009 collection of 836,135 documents captured from Enron by the Federal Energy Regulatory Commission. During the course of the TREC 2009 Legal Track Interactive Task, these documents were reviewed for responsiveness to seven “topics” (requests for production composed by TREC) by several participating teams using various technology-assisted review methods, which are detailed in the TREC proceedings.⁴⁹ Two teams showed superior effectiveness over five of the topics, according to F_1 (the harmonic mean of recall and precision). Grossman and Cormack compared these results to those of a human review of a statistical sample conducted (for three topics) by professional contract reviewers and (for two topics) by third-year law students. For all topics and all measures (recall, precision, and F_1) either the technology-assisted review was superior, or the measured difference was not statistically significant. On average, the technology-assisted reviews achieved recall, precision, and F_1 scores of 76.7%, 84.7%, and 80.0%, respectively, while the human reviews achieved 59.3%, 31.7%, and 36.0%, respectively. Overall, these studies indicate that one should not presume human review to be the most effective approach, that certain technology-assisted review methods can improve on human review, and that no review method is perfect. On the other hand, these studies do not indicate that all technology-assisted review methods are more effective than human review in all circumstances.

The limitations of keyword approaches to search and retrieval first exposed in the Blair and Maron study, and validated in subsequent research, have not faulted the ability of computers to locate documents meeting the attorneys’ search criteria – but rather the inability of the attorneys and paralegals to anticipate all of the possible ways that people might refer to the issues in the case. The richness and ambiguity of human language causes severe challenges in identifying relevant information.

As Blair and Maron (and subsequent studies) demonstrate, human language is highly ambiguous and full of variation. In the years since Blair and Maron, and with increasing attention focused on the e-discovery space, the IR community has been engaged in research and the development of methods, tools, and techniques that compensate for endemic ambiguity and variation in human language, thereby improving the recall and precision of searches.

documents found by all combined search methods); see generally, *TREC Legal Track Overview Papers, 2006-2011*, <http://trec-legal.umiacs.umd.edu/>.

48 See Maura R. Grossman & Gordon Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J. L. & TECH. 11 (2011), available at <http://jolt.richmond.edu/v17i3/article11.pdf>; Herbert L. Roitblat, et al., *Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review*, 61 J. AM. SOC’Y FOR INFO. SCI. & TECH. 70 (2010), available at <http://jolt.richmond.edu/v17i3/article11.pdf>; Patrick Oot, et al., *Mandating Reasonableness In A Reasonable Inquiry*, 87 DENV. U. L. REV. 533 (2011), available at http://law.du.edu/documents/denver-university-law-review/v87-2/Oot_PDF.pdf; see also Howard Turtle, *Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance*, 1994 PROCEEDINGS OF THE 17TH ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL 212220 (using structured case law in Westlaw databases), see generally, RAND 2012 STUDY, pp. 59-69 (summarizing results from these and other recent studies).

V. BOOLEAN AND BEYOND: A WORLD OF SEARCH METHODS, TOOLS, AND TECHNIQUES

In the decades since the Blair and Maron study, a variety of new search tools and techniques have been introduced to help find relevant information and weed out irrelevant information. Understanding these various tools and methods is critical. All automated methods are not created equal and do not perform the same functions and tasks. It is important to know what each methodology does when it is used alone or in conjunction with other methods and which tools are most effective for which purposes.

Clearly, different search methods have different functions and values in different circumstances. There is no one best system for all situations, an important fact for practitioners learning the techniques of search and retrieval technology to understand.

A more detailed description of search methods and techniques is set out in the Appendix. These methods can be grouped into four broad categories, but there are hybrid and crosscutting approaches that defy easy placement in any particular “box.”

Keywords and Boolean Operators

First, there are keyword-based methods, ranging from the simple use of keywords alone, to the use of strings of keywords with what are known as “Boolean operators” (including AND, OR, “AND NOT,” or “BUT NOT”).

Categorizations of Data Sets

Second, there are other techniques relying on categorizations of the entire data set with various methodologies heavily reliant on deriving (i.e., coming to a consensus on) a thesaurus, taxonomy, or “ontology” of related words or terms. These techniques can be used to categorize the entire data set into specified categories all at once or as more data is added to the data set.

However, data sets generally need to be indexed to use any of the latter methodologies – where the indexing will take more time depending on what one indexes (e.g., indexing all of the data will take substantially longer than indexing selected fields).

There are a variety of indexing tools, some of which are available as open source tools. Indexing structured data may take less time than indexing data in an unstructured form, if only designated fields are indexed. Indexing an unstructured data set is time consuming because of the need to index all the words (except for “and,” “a,” “the,” or other common “noise” words). Knowing what is being indexed will be important to set expectations in terms of timing and making the data useful for querying or review.

Alternative search methods to keywords can, in some instances, free the user from having to guess, for every document, what word the author might have used. For example, there are more than 120 words that could be used in place of the word “think” (e.g., guess, surmise, anticipate and so on). As the Blair and Maron study shows, people are actually very poor at guessing the right words to use as search terms to find the documents they are looking for without overwhelming the retrieval with irrelevant documents. In light of this fact, alternative search methods help organize large collections of documents which humans would otherwise be unable to organize.

Using a thesaurus, taxonomy, or ontology generally provides the results one would expect, because these systems explicitly incorporate one's expectations about what is related to what. They are most useful when one has (or can buy) a good idea of the conceptual relations to be found in one's documents – or one has the time and resources needed to develop them. Clustering, Bayesian classifiers, and other types of systems have the power to discover potentially unanticipated relationships in the text. This means that one gets unexpected results from time to time, which can be of great value, but can also be somewhat over-inclusive (or even wrong). An example: after training on a collection of medical documents, one of these systems learned that Elavil and Klonopin were related (they are both anti-anxiety drugs). A search for Elavil turned up all the documents that contained that word, along with “false positive” documents containing only the word “Klonopin” (without the word “Elavil”).

Such systems can discover the meaning of at least some acronyms, jargon, and code words appropriate to the context of the specific document collection. No one has to anticipate their usage in all possible contexts; the systems, however, can help to derive them directly from the documents.

Computer- or Technology-Assisted Review Methods

Third, there are a variety of methods that combine both technological and human inputs in an iterative search design. These techniques stem from various research directions in Artificial Intelligence and can be categorized under the general rubric of computer- or technology-assisted review. When drawing on machine-learning techniques, these tools are sometimes referred to as Predictive Coding⁴⁹; however, the more inclusive nomenclature for the entire spectrum of advanced methods remains computer- or technology-assisted review. Generally put, computer- or technology-assisted approaches are based on iterative processes where one (or more) attorneys or IR experts train the software, using document exemplars, to differentiate between relevant and non-relevant documents. In most cases, these technologies are combined with statistical and quality assurance features that assess the quality of the results. The research cited above has demonstrated such techniques superior, in most cases, to traditional keyword based search, and, even, in some cases, to human review.⁵⁰

The computer- or technology-assisted review paradigm is the joint product of human expertise (usually an attorney or IR expert working in concert with case attorneys) and technology. The quality of the application's output, which is an assessment or ranking of the relevance of each document in the collection, is highly dependent on the quality of the input, that is, the human training. Best practices focus on the utilization of informed, experienced, and reliable individuals training the system. These individuals work in close consultation with the legal team handling the matter, for engineering the application. Similarly, as explained below, the defensibility and usability of computer- or technology-assisted review tools require the application of sound approaches to selection of a “seed” or “training” set of documents, monitoring of the training process, sampling, and quantification and verification of the results.⁵¹

Well-thought-out techniques are needed in order to ensure that the set of documents used to train the system provides thorough coverage of the entire document

49 See Bruce Hedin, et al., *Overview of the TREC 2009 Legal Track*, <http://trec.nist.gov/pubs/trec18/papers/LEGAL09.OVERVIEW.pdf> (last visited November 25, 2013).

50 See *supra*, notes 48-49.

51 See generally, *The Sedona Conference Commentary on Achieving Quality in E-Discovery*, *supra* note 33.

population and particularly, of the relevant material (both expected and unexpected) contained therein. Sampling will be an essential component of any effective quality control regimen; it is important that samples drawn for quality control are separate from those used for training, in order to ensure independence of statistical measurement.

Iterative monitoring of the training process ensures that the system is adequately trained, (meaning that additional exemplars could not substantially enhance review effectiveness), while avoiding wasteful use of expensive “expert” resources through excessive training. In some cases, active learning techniques are applied to accelerate the training process, reducing to a minimum the number of training documents required, and avoiding the inefficiencies of random sampling, especially in low prevalence populations. Active learning systems select new exemplars for training based on knowledge of the population that the application has generated from previous training examples.

Statistically valid *measurement techniques*, based on precision and recall, as described above, are expressed within confidence intervals and at certain confidence levels to estimate results. Quality assurance techniques are critical in order to verify outcomes; for example, sampling of the documents culled as irrelevant to verify that they contain the expected low prevalence of relevant documents.

None of these systems is magical. Language is sometimes shared between two people who invent a shorthand or code. And all tools require a healthy dose of good legal judgment, based on a well thought out approach. Some techniques may be difficult to understand to those without technical backgrounds, but they need not be mysterious. If a vendor cannot explain how a system works, then the buyer should beware and either require an explanation or consider an alternate approach.

There is no magic to the science of search and retrieval; it is comprised of mathematics, linguistics, computer science, and hard work. If lawyers do not become conversant in this area, they risk surrendering the field’s intellectual jurisdiction to other disciplines, as well as risk poor quality and costly e-discovery outcomes.

VI. PRACTICAL GUIDANCE FOR EVALUATING THE USE OF AUTOMATED SEARCH AND RETRIEVAL METHODS

Practice Point 1. ***In many settings involving large amounts of relevant electronically stored information (ESI), relying solely on a manual search process for the purpose of finding responsive documents may be infeasible or unwarranted. In such cases, the use of automated search methods should be viewed as reasonable, valuable, and even necessary under certain circumstances.***

For the reasons articulated in prior sections, the demands placed on practitioners and parties in litigation and elsewhere increasingly dictate that practitioners must evaluate the use of automated search and retrieval methods in a wide variety of cases and contexts. Particularly (but not exclusively) in large and complex litigation, where discovery is expected to encompass hundreds of thousands to hundreds of millions of potentially responsive electronic records, there is no reasonable possibility of marshalling the human labor required to undertake a document-by-document, manual review of the potential universe of discoverable materials. This is increasingly true both for parties responding to a discovery request and for parties who propound discovery (and receive a massive amount of

material in response). Where the infeasibility of undertaking manual review is acknowledged, utilizing automated search methods may not only be reasonable and valuable, but also necessary.

Even in less complex settings, overreliance on manual review may be an inefficient use of scarce resources. This is especially the case where automated search tools used on the front end of discovery could prove useful in a variety of ways, including early case assessment for settlement or other purposes, or for prioritizing or grouping documents to allocate resources or facilitate later manual review.

Of course, the use of automated search methods is not intended to entirely eliminate the need for manual review; indeed, in many cases, both automated and manual searches will be conducted, with initial automated searches used for culling down a universe of material to more manageable size (or prioritizing documents), followed by a secondary manual review process. So too, while automated search methods may help identify privileged documents within a larger set, the majority of practitioners may still rely on largely manual review processes to identify the basis for the assertion of privilege.

Practice Point 2. The successful use of any automated search method or technology will be enhanced by a well-thought-out process with substantial human input on the front end.

As discussed above, the decision to employ an automated search method or technology cannot be made in a vacuum, on the assumption that the latest “tool” will solve an attorney’s discovery obligations. Rather, to maximize the chances of success in terms of finding responsive documents, a well-thought-out strategy capitalizing on “human knowledge” available to a party should be implemented at the earliest opportunity. This knowledge can take many forms.

First, a party must evaluate the specific legal setting, since the nature of the lawsuit or investigation, the field of law involved, and the specific causes of action under which a discovery obligation may arise must all be taken into account. For example, keyword searches alone in highly technical patent cases may prove highly efficacious. But in other types of cases, including those with broad causes of action involving subjective states of intent, a practitioner should consider alternative search methods.

Second, in any legal setting involving consideration of automated methods for conducting searches, counsel and client should perform an analysis to define the target universe of documents that is central to the relevant causes of action. This would include not only assessing relevant subject areas and “drilling down” with as much specificity as possible, but also analyzing the custodians (or others) who would be in possession of such relevant data. Time and cost considerations must also be factored in, including budgeting for human review time. These practice points apply whether a party is a defendant and holds a universe of potentially discoverable data, or a plaintiff who is expecting to receive a massive data set in response to requests for production.

Practice Point 3. The choice of a specific search and retrieval method will be highly dependent on the specific legal context in which it is to be employed. Parties and their counsel must match the use case with the tools and best practices appropriate to address it and must incorporate proportionality considerations involving the overall cost the stakes and the stakes of the litigation.

The choice of a search and retrieval method for a given situation depends upon a number of factors. Two of the biggest decisions to make are the acceptable level of false positive “noise” (i.e., achieving higher “precision”), and the acceptable level of false negatives (i.e., maximizing “recall”). There are a number of overarching factors that lawyers should consider in evaluating the use of particular search and retrieval methods in particular settings.

First, the “heterogeneity” of the relevant universe of ESI is a significant factor. ESI that is potentially relevant may be found in multiple locations, and in a variety of forms, including structured and unstructured active computer environments, removable media, backup tapes, and a variety of email applications and file formats. In some cases, information that provides historical, contextual, tracking, or managerial insight (such as metadata) may be relevant to a specific matter and demand specialized data mining search tools. But in other cases, the very same data will be irrelevant.

Second, the volume, prevalence, and condition of the likely relevant ESI, and the extent to which ESI is contained within static or dynamic electronic applications, are all relevant to the party’s or its counsel’s decisions.

Third, for any particular search and IR method, the time it takes and its cost (compared to other automated methods or human review) must also be considered.

Fourth, the goals of the search are a factor (e.g., capturing or finding as many responsive documents as possible regardless of time and cost versus finding responsive documents as efficiently as possible, i.e., with the least number of non-responsive documents). In other words, the practitioner must consider the desired trade-off between precision and recall. Given the particular setting, the party seeking to employ one or more search methods should assess the relative importance in that setting of finding responsive ESI versus the importance of eliminating non-responsive data. Depending on this assessment, one or more alternative search methodologies may prove to be a better match in the context of a particular task.

Fifth, one must consider the skills, experience, financial, and practical constraints of the representatives of the party making the selection (e.g., the attorneys, litigation support staff, vendors, the Special Master, etc.).

Sixth, the current status of electronic discovery in the matter, including the extent to which activities including preservation and collection are occurring in addition to processing and/or attorney review.

Seventh, one must investigate empirical research supporting the reliability of the search and IR method for particular types of data, or in particular settings.

* * *

Although not the focus of this Commentary, practitioners may also wish to consider the use of advanced search methods at earlier stages of the e-discovery process before traditional document review.

- **Early Case Assessment:** A key objective of early case assessment is to assess case risk and cost, with the goal of avoiding futile and wasteful litigation activity. Tools with sophisticated metadata search and computer- or technology-assisted

review capabilities can be helpful in identifying a small subset of highly probative documents within the population with high precision. Users can then focus on this “rich” subset of data to make rapid and informed decisions on case strategy.

- **Culling:** For proper use in culling strategies, search and retrieval technologies should support a valid statistical framework able to estimate characteristics such as prevalence, precision, and recall. This facilitates a user’s quantifiable assessment of the cost and risk involved in culling decisions: for example, whether culling a certain set of documents will potentially yield 5% or 95% of the relevant documents.
- **Prioritized review:** Prioritized review structures the review process to start with the documents most likely to be relevant, progressing to the documents least likely to be relevant. Stratified review, a variation of prioritized review, matches document importance and reviewer quality, such that more skilled reviewers can review documents requiring greater expertise. Prioritized and stratified reviews require computer- or technology-assisted review tools that are able to classify documents as “more likely to be relevant” through “less likely to be relevant.”

Finally, in adopting proportionality, parties need to balance the costs of using alternative search methods with the perceived benefits and risks in a given litigation context. Costs and time will vary depending on the desired rate of recall (how many documents need to be found) and precision (how many documents need to be reviewed to yield relevant documents). Risk also is reflected in recall, i.e., how many documents are found versus how many are “left on the table.” The proportionate costs versus benefits and risks that a user is willing to bear are a function of what technology is reasonably available and what is at stake in the particular matter, taking into account what e-discovery phase is being addressed (document review versus other aspects of discovery, at earlier stages, per the above).

Practice Point 4. Parties and their counsel should perform due diligence when choosing a particular information retrieval product or vendor service.

The prudent practitioner should ask questions regarding search and retrieval features and the specific processing and searching rules that are applied to such features. Some tools are fully integrated into a vendor’s search and review system, whereas others are “stand alone” tools that may be used separately from the particular review platform. It is essential not only to understand how the various tools function, but also to understand how the tools fit within the overall workflow planned for discovery. A practitioner should inquire as to which category or categories the specific tool fits into, how it functions, and what third-party technology lies behind the tool.

It is also essential that specific methods or tools be made understandable to the court, opposing parties, and the attorney’s client. How data is captured and indexed (and how long it takes to build an index) also may affect a decision on use; it is therefore important to understand how a particular system deals with rolling input and output over time, considering its flexibility and scalability. The ability to perform searches across metadata, to search across multiple indices or stores of data, to search embedded data, to refine search results (nested searches), to save queries, to capture duplicates and perform

deduplication, to trace email threads, and to provide listings of related terms or synonyms are all examples of specific functional requirements that should be clarified depending on case needs.

Other types of due diligence may involve administrative matters (e.g., understanding maintenance and upkeep, additional charges, system upgrades, availability of consultants or technicians to address problems, system performance), quality control issues (e.g., prior testing of the method or tool in question; how databases and dictionaries supporting concept search were populated; the strength of the provider's application development group), and, finally, any relevant licensing issues which could involve proprietary software or escrow agreements with third parties.

Practice Point 5. Because of the characteristics of human language, no search and information retrieval tool can guarantee the identification of all responsive documents in large data collections. Moreover, different search methods may produce different results, subject to a measure of statistical variation inherent in the science of information retrieval.

Just as with past practice involving manual search through traditional paper document collections, there is no requirement that “perfect” searches will occur – only that parties and their counsel act reasonably and in good faith in the performance of their discovery obligations. From decades of IR research, it is clear that a 100% rate of recall, i.e., the ability to retrieve all responsive documents from a given universe of electronic data, is an unachievable goal. As discussed in prior sections, the richness of human language, with its attendant elasticity, causes all present day automated search methods to fall short of perfection.

Moreover, there will always be a measure of statistical variation associated with alternative search methods, i.e., some responsive documents will be found by one search method while being missed by others. Even the same search method may return different results if new documents are added to the searched universe. Particularly in the context of a large data set, a search method should be judged by its overall results (such as using measures of recall and precision), rather than being judged by whether it produces the identical document set as compared with a different technique.

However, it is important not to compare “apples with oranges.” Given the present state of information science, it would be a mistake to assume that one search method will work optimally across all types of possible inquiries or data sets (e.g., what works well in finding word processing documents in a given proprietary format may not be as optimal for finding information in structured databases, or in a collection of scanned images). This is another area where, consistent with the above principles, a good deal of thought should be given at the outset to the precise problem, in terms of its scope and relevancy considerations, before committing to a particular search method.

Practice Point 6. Parties and their counsel should make a good faith attempt to cooperate when determining the use of particular search and information retrieval methods, tools, and protocols (including keywords, concepts, computer- and technology-assisted review and other types of search parameters and quality control measures).

The body of case law that has emerged since 2006⁵² indicates that courts are becoming more comfortable with addressing search and retrieval issues, particularly in the context of approving protocols, or ordering parties to share information that would lead to the development of more refined search protocols. The fact that some courts have waded into these issues demonstrates how rapidly the law has been evolving since the 2006 amendments to the Federal Rules of Civil Procedure.⁵³

Under Rule 26(f), the parties' initial planning should address "[a]ny issues relating to disclosure or discovery of electronically stored information," as well as "[a]ny issues relating to preserving discoverable information." These initial discussions on preservation and production should include a specific discussion on search methods and protocols to be employed by one or both parties.⁵⁴ While disclosure of these methods and protocols is not mandated or legally required under this rule, the advantages of collaborating should strongly be considered. In many cases, reaching an early consensus on the scope of searches can minimize the overall time, cost, and resources spent on such efforts, as well as minimize the risk of collateral litigation challenging the reasonableness of the search method employed.⁵⁵

The Sedona Conference Cooperation Proclamation, published in 2008 underscores Practice Point 6, here, by including among the methods for accomplishing cooperation in e-discovery: (i) "Exchanging information on relevant data sources, including those not being searched;" and (ii) "Jointly developing automated search and retrieval methodologies to cull relevant information."⁵⁶

Practice Point 7. Parties and their counsel should expect that their choice of search methodology (and any validation of it) will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and at trial).

Counsel should be prepared to explain what keywords, search protocols, and alternative search methods were used to generate their production set, including ESI made subject to a subsequent manual search for responsiveness and privilege. This explanation may need to come from a technical "IT" expert, a statistician, or an expert in search and retrieval technology. Parties should anticipate that in contested matters, an opposing party may request the justification of particular search methods used; this may require a demonstration of the recall and precision (or other measures) for the output of a chosen search method.⁵⁷ Counsel must be prepared to answer questions and even prove the reasonableness and good faith of their methods.

52 See Baron, *Law in the Age of Exabytes*, *supra* note 24.

53 See Kenneth J. Withers, *Electronically Stored Information: The December 2006 Amendments to the Federal Rules of Civil Procedure*, 4 NW. J. OF TECH. & INTELL. PROP. 171 (2006) (what "probably strikes the reader [of Treppel], *supra* note 26] as matter of fact, sensible, and routine, would have been extraordinary a scant six years ago, when the last major revision of the discovery rules went into effect [in 2000])."

54 See, e.g., *In re Pilot Project Regarding Case Management Techniques for Complex Civil Cases in the Southern District of New York* (Standing Order, Nov 1, 2011), http://www.nysd.uscourts.gov/cases/show.php?db=notice_bar&cid=261.

55 See Paul and Baron, *Information Inflation*, *supra* note 14, at paras. 50-55 (discussing an iterative collaboration process that includes adoption of multiple "meet and confers" to discuss and refine preliminary search results).

56 *The Sedona Conference Cooperation Proclamation*, available at <https://thesedonaconference.org/download-pub/3802>. A companion piece, *The Case for Cooperation*, published in *The Sedona Conference Journal*⁸ goes on to say:

[W]orking cooperatively with opposing counsel to identify a reasonable search protocol, rather than making boilerplate objections to the breadth of a requested protocol or unilaterally selecting the keywords used without disclosure to opposing counsel, may help avoid sanctions or allegations of intentional suppression. Indeed, because knowledge of the producing party's data is usually asymmetrical, it is possible that refusing to 'aid' opposing counsel in designing an appropriate search protocol that the party holding the data knows will produce responsive documents could be tantamount to concealing relevant evidence.

10 SEDONA CONF. J. 339, 344 (2009 Supp.).

57 See, e.g., *Moore*, 287 F.R.D. at 182; *Kleen Prod.*, 2012 WL 4498465 at *1.

Practice Point 8. Parties, counsel, and the courts should be alert to new and rapidly evolving search and information retrieval methods. Moreover, parties and their counsel should recognize that Information Retrieval is a distinct field of study that includes expertise in such areas as computer science, statistics, and linguistics, and that consultation with or utilization of experts in information retrieval may improve the quality of search results in complex cases involving large volumes of ESI.

Given the rapid evolution of technology, what constitutes a reasonable search and IR method is subject to change. The legal community needs to be vigilant and must examine new and emerging techniques and methods which can yield better search results. In particular settings, lawyers should endeavor to incorporate evolving technological progress at the earliest opportunity in the planning stages of discovery or other legal setting involving search and retrieval issues.

Successful search of large amounts of ESI is increasingly dependent on the expertise of those doing the searches. Attorneys who lack expertise in IR or statistics should consider consulting or collaborating with IR experts, when appropriate, in complex cases. In general, the Bar would do well to understand that a greater appreciation of IR and scientific methods will result in better overall search. This is analogous to the situation where a tool, such as a scalpel, could in theory be used by anyone, regardless of their expertise in medicine or surgery. But when that tool is used in the hands of a trained surgeon (as opposed to someone who lacks that expertise), common sense dictates that the results will likely be better.

VII. FUTURE DIRECTIONS IN SEARCH AND RETRIEVAL SCIENCE

What prospects exist for improving present day search and retrieval methodologies? And how can lawyers play a greater role in working with the IR research community to improve the accuracy and efficiency of search and review technology?

A. Harnessing the Power of Artificial Intelligence (AI)

A statement from page 36 of The Sedona Conference, *Navigating The Vendor Proposal Process* (2007 ed.), under the general heading “Advanced Search and Retrieval Technology,” bears repetition here: “Technology is developing that will allow for electronic relevancy assessments and subject matter, or issue coding. These technologies have the potential to dramatically change the way electronic discovery is handled in litigation, and could save litigants millions of dollars in document review costs. Hand-in-hand with electronic relevancy assessment and issue coding, it is anticipated that advanced searching and retrieval technologies may allow for targeted collections and productions, thus reducing the volume of information involved in the discovery process.”

The growing enormity of data stores, the inherent elasticity of human language, and the unfulfilled goal of computational thinking to approximate the ability and subtlety of human language all present steep challenges to the IR and AI communities in their quest to develop optimal search and retrieval techniques.

But the future holds promise. Not only is there available technology to apply sophisticated computer algorithms to data mine traditional text, but new and better

approaches to image and voice pattern recognition are looming on the horizon. Indeed, there is already rudimentary technology available for searching audio by search terms, and some processes are confronting – and succeeding – in searching by image or picture.⁵⁸ Clearly, at some point, all forms of data stored in corporations and institutions will be within the scope of future information demands in legal settings, and likely within the ambit of future automated searching processes.

Finding information on the Web sometimes is easier than finding documents on one's own hard drive. The post-Google interest in building better search engines for the Web can only lead to new and better search techniques applied to more well-defined contexts, such as corporate and institutional intranets and data stores.

A “2020 Science” report issued by Microsoft in March 2006 anticipated the near-term development of “novel data mining technologies and novel analysis techniques,” including “active learning” in the form of “autonomous experimentation” and “artificial scientists,” in replacement of “traditional machine learning techniques [that] have failed to bring back the knowledge out of the data.”⁵⁹ With the emergence of computer- or technology-assisted review methods in e-discovery, we are beginning to see such techniques played out in litigation. Beyond the short-term horizon, scientists are expected to embrace emergent technologies including the use of genetic algorithms, nanotechnology, quantum computing, and a host of other advanced means of information processing. And future AI research in the specific domain of search and retrieval is unbounded and, at least in part, unpredictable.

B. The Role of Process in the Search and Retrieval Challenge

Every search and retrieval technology has its own methodology to ensure the technology works properly, relying on a set of instructions that outline the workflow for the tool. How well these methods are applied significantly impacts the performance, and therefore the results achieved by the technology. This is where process comes in to play. Process provides order and structure by setting guidelines and procedures designed to ensure that a technology performs as intended. Effectively applied, process drives the consistent and predictable application of the search and retrieval technology. The results derived from the consistent and predictable application of search and retrieval tools establish the technology's credibility and value.

The Importance of Process

A process is a considered series of events, acts, or operations leading to a predictable result or effect. A process, like a technology, is a “tool” that can be used to assist in completing a task. The use of a well-defined and controlled process promotes consistency, reliability, and predictability of the results and ensures the efficient use of the resources required to produce them. As such, a process does not find the answer to or attain the objective of a task on its own. Process, no matter how well designed and executed, cannot replace the exercise of judgment; however, process promotes the exercise of judgment by ensuring that the most accurate and reliable information is available when making decisions. In the search and retrieval context, this means the availability of consistent and reliable information to assist the parties in making informed decisions.

58 See, e.g., Idée, Inc., *TinEye: Reverse Image Search*, <http://www.tineye.com/> (last visited November 25, 2013).

59 See Microsoft, *2020 Science*, at 15, available at <http://research.microsoft.com/towards2020science/downloads.htm>.

The use of process promotes consistency by establishing a defined approach to a task. The resulting consistency promotes reliability and predictability. Reliability and predictability allow for better planning, performance, and cost management. Altogether, risk is reduced and confidence is promoted.

One can visualize search and retrieval as a process enabling a party to distinguish potentially discoverable information from among a broader set of electronic data collected for the purposes of production. It consists of several steps that take place in the context of a particular search and retrieval technology. Because the application of process is flexible, it can be used to address unique conditions that might be associated with a technology, such as where the use of a search and retrieval technology itself creates issues. For example, the use of search and retrieval technologies to address significant volumes of information may not address all problems: as review volumes increase (even with carefully crafted and tested search criteria) the likelihood of being swamped by false positives or missing false negatives increases greatly. By developing and implementing process steps which consistently address these issues, their impact can be diminished and the reasonableness and good faith of the technology can be established.

“Process” as a Measure of Reasonableness and Good Faith

Search and retrieval in this new era requires the establishment and recognition of a new standard. A standard of absolute perfection is and always has been unrealistic; but now, with quantitative data available, we know perfection is not only unrealistic but also quite simply unachievable.

Rather than perfection, which would demand the identification and production of every relevant, non-privileged document, the standard against which to measure these new technologies and processes should incorporate the same principles that have traditionally governed all discovery: reasonableness, good faith, and proportionality.⁶⁰ Although these terms raise concerns about ambiguity and uncertainty, they can actually represent a well-defined set of expectations in the context of the discovery process.

A process that emphasizes reasonableness, good faith, and proportionality is fully consistent with what is required under the discovery process. Discovery of information relevant to a dispute gathered by an opponent is often central to a fair and efficient resolution.⁶¹ A party need only identify and produce that which is relevant, as defined by the rules, with the degree of diligence expected and available by experienced practitioners acting reasonably.⁶² As noted in Sedona Principles 6 and 11, a party may choose to implement this approach in a reasonable manner.

Sound process applied to the use of search and retrieval technology can readily establish a measurable means for conducting discovery that satisfies the Rules. Reasonableness, good faith, and proportionality can be defined and measured by identifying performance criteria based on their attributes. Accordingly, the unreasonable and unattainable goal of “perfection” should not hinder the attainable and measurable goal of reasonableness.

60 See Fed. R. Civ. P. 26(g) & 26(b)(2)(C).

61 See *Hickman*, 329 U.S. at 507.

62 Under Fed. R. Civ. P. Rule 26(g)(1), an attorney of record is expected to certify that to the best of his or her “knowledge, information, and belief, formed after a reasonable inquiry,” that disclosures are “complete and correct” as of the time they were made. Similarly, under Rule 26(g)(2), an attorney must certify that to the best of his or her “knowledge, information, and belief, formed after a reasonable inquiry,” that discovery “requests, responses, and objections” are made “consistent with these rules.” See also Fed. R. Civ. P. 26(b)(2)(C).

As search and retrieval technologies and associated processes are developed, parties will no doubt want to use them to achieve defensible and credible results. If a party fails to adhere to appropriate performance guidelines, it will be subject to scrutiny and criticism, and perhaps even sanctions. Therefore, an established process – in conjunction with sound technology – can serve as a benchmark for conducting future discovery. Further, defensibility with opposing counsel and the court will likely depend implementing and adhering to processes developed for use with search and retrieval technologies.

Implementing Process

Using a search and retrieval technology in conjunction with an implementing process will involve iterative activity. This will incorporate feedback loops at appropriate decision points to allow integration of what a case team learns after each step of the process. This, in turn, will calibrate and maximize the technology's ability to identify relevant information. It is through this feedback that case teams will acquire sound information to use in making both strategic and tactical decisions.

The initial search and retrieval process should be designed as a “pilot” process that can be evaluated and modified as the team learns more about the corpus of information to be reviewed. One useful approach initiates the process by focusing on the information collected from a few key custodians at the center of the facts at issue in the litigation or investigation. Focusing on information collected from core custodians (information with a higher likelihood of relevance) will help the team efficiently understand the issues and language used by the custodians, and enable them to more efficiently develop and implement an appropriate search and retrieval process for subsequent custodians and ESI.

The initial selection and refinement of search terms can also benefit from sampling techniques used to rank the effectiveness of various terms or concepts. Reviewing samples of information that include selected search terms or concepts and ranking their relative value based on their efficacy in retrieving relevant information (recall) and their efficiency in excluding non-relevant information (precision) can help focus the selection of appropriate search terms.⁶³

The development of process control logs and improved second-level review techniques can also help the review team consistently apply the designed process to all of the information to be reviewed. Additionally, a second-level review process based on statistical sampling techniques can ensure acceptable levels of quality. While these techniques are relatively unknown in the typical review processes in use today, their widespread adoption in businesses of all types should drive their implementation in large document review projects in the near future.⁶⁴

C. How The Legal Community Can Contribute to The Growth of Knowledge

Human review of documents in discovery is expensive, time consuming, and error-prone. The application of linguistic and statistically-based content analysis, search and retrieval technologies such as computer- or technology-assisted review, and other tools, techniques, and process in support of the review function can effectively reduce cost, time, and error rates.⁶⁵

⁶³ See *supra*, text at Part IV.

⁶⁴ See generally The Sedona Conference *Commentary on Achieving Quality in E-Discovery*, *supra* note 33.

⁶⁵ See *supra* note 48 and accompanying text.

Recommendations

1. *The legal community should support collaborative research with the scientific and academic sectors aimed at establishing the efficacy and efficiency of a range of automated search and information retrieval methods.*
2. *The legal community should encourage the establishment of objective benchmarking criteria, for use in assisting lawyers in evaluating the competitive legal and regulatory search and retrieval services market.*

Up until recently, in the years since the 1985 Blair and Maron study, there was little in the way of peer-reviewed research establishing the efficacy of various methods of automated content analysis, search, and retrieval as applied to a legal discovery context. Research into the relative efficacy of search and retrieval methods should acknowledge that each alternative should be viewed in the context of its suitability to specific document review tasks. Different technologies, tools, and techniques obviously have different strengths and weaknesses. Moreover, the outcomes of the application of advanced content analysis, search, and retrieval methods may be significantly different based on expertise of the operator. Ideally, research should advance the goal of setting a minimum or baseline standard for what constitutes an adequate information retrieval process, as well as reaching agreement on how to benchmark competing methods against agreed-upon objective evaluation measures.

Since 2006, The Sedona Conference has supported the TREC Legal Track (part of the TREC research program run by the National Institute of Standards and Technology). NIST is a federal agency that collaborates with industry and academia to develop and apply technology, measurements, and standards. TREC is designed “to encourage research in information retrieval from large text collections.”⁶⁶ The TREC legal track has involved evaluation of a set of search methodologies based on lawyer relevancy assessments on topics drawn from large publicly available document databases. The results that have come out of the TREC Legal Track represent the type of objective research into the relative efficacy of Boolean and other search methods that the legal community should further encourage.⁶⁷

However, a need exists to expand upon TREC research to accommodate the potential retrieval of tens or hundreds of millions of arguably relevant documents among a greater universe of terabytes, petabytes, exabytes, and beyond, and to study new and emerging forms of ESI, including text messaging and all forms of online social media. Members of The Sedona Conference community have and will continue to participate in collaborative workshops and other fora focused on issues involving IR.⁶⁸ How best to

⁶⁶ The Text Retrieval Conference (TREC) was started in 1992. Its purpose is to support research within the IR community by providing the infrastructure necessary for large-scale evaluation of text retrieval methodologies. TREC is overseen by a program committee consisting of representatives from government, industry, and academia. Each TREC track involves a test database of documents and topics. Participants run their own search and retrieval systems on the data, and return to NIST a list of the retrieved top-ranked documents. NIST generally pools the individual results, judges the retrieved documents for correctness, and evaluates the results. The TREC cycle ends with a workshop forum for participants to share their experiences. The TREC test collections and evaluation software are available to the IR research community at large, so organizations can evaluate their own retrieval systems at any time. TREC has successfully met its dual goals of improving the state-of-the-art in IR and of facilitating technology transfer, and some of today’s commercial search engines include technology first developed at TREC. For further information, see <http://trec.nist.gov>.

⁶⁷ See TREC Legal Track Overview Papers, 2006-2011, <http://trec-legal.umiacs.umd.edu/>.

⁶⁸ See, e.g., Discovery of Electronically Stored Information (“DESI”) Workshops, held at the Eleventh, Twelfth, Thirteenth, and Fourteenth International Conferences on Artificial Intelligence and Law (ICAIL 2007, ICAIL 2009, ICAIL 2011, ICAIL 2013) (links available at <http://www.umiacs.umd.edu/~oard/desi5/>); Information Retrieval for E-Discovery (SIRE) Workshop, SIGIR 2011 (link available at <http://www.umiacs.umd.edu/~oard/sire11/>).

leverage the work of the IR community to date is an enterprise beyond the scope of this paper. The Sedona Conference intends to remain in the forefront of the efforts of the legal community in seeking out centers of excellence in this area, including the possibility of fostering private-public partnerships aimed at focused research.

APPENDIX: TYPES OF SEARCH METHODS

This Appendix is a “survey” of some of the different search methods found in the information retrieval literature, which form the basis of offerings by vendors in the legal marketplace. The list is not exhaustive. Indeed, as the main body of the Commentary makes clear, rapid technological progress will inevitably affect how methods are described, implemented, and subsequently replaced with new ways of performing search and retrieval.

Three further notes on this survey are in order. First, the following search methods are not intended to be mutually exclusive. Indeed, many products encourage the use of hybrid, combined, or cumulative approaches to search.

Second, the choice of method is condition- and objective-specific. The method best suited to one circumstance may not be the best suited to another circumstance. The methods reviewed in this survey may be viewed as distinct tools in the search “toolbox,” each with its appropriate application or applications.

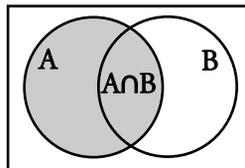
Third, potential users of the technologies reviewed in this survey are reminded that the tools alone cannot guarantee an effective search, any more than a well-made scalpel guarantees a successful surgery. Search tools will be effective only if applied with sound methodological principles and with appropriate expertise.

A. Boolean Search

A “Boolean search” utilizes the principles of Boolean logic named for George Boole, a British born mathematician. Boolean logic is a method for describing a “set” of objects or ideas. Boolean logic was applied to IR as computers became more widely accepted. Boolean searches can easily be applied to large sets of unstructured data and return results which exactly match the search terms and logical connectors applied by the operators.

As used in set theory, a Boolean notation demonstrates the relationship between sets or groups of information – in our example, two sets of information, “A” and “B” – and, in effect, creates a new set of information.

If a search seeks information contained within either original set “A” *or* original set “B” (essentially any area within either set in the Venn diagram below), the searcher is creating a “Union” of A and B (denoted “ $A \cup B$ ”). If the search seeks information which would be found within both set “A” *and* set “B,” if each set was searched separately (the area within the overlap of the Venn diagram below), the searcher is creating an “Intersection” of A and B (denoted “ $A \cap B$ ”). A Venn diagram picture easily depicts these relationships (see below).



The “OR” Boolean operator directs that the set may contain any, some, or all of the keywords searched. The purpose of this command is to encompass alternative

vocabulary terms. **OR** is represented by the union of the sets (“A∪B”) (the entire area within both the circles above). The use of **OR** *expands* the resulting Boolean set.

The “**AND**” Boolean operator identifies the intersection of two sets or two keywords. The purpose of this command is to help construct more complex concepts from more simple vocabulary word “building blocks.” **AND** is represented by intersection of the sets (“A∩B”) (the shaded area within the intersection of the two circles). The use of **AND** *restricts* the resulting Boolean set.

The “**NOT**” Boolean operator eliminates unwanted terms. The purpose of this command (often preceded by either “**AND**” or “**BUT**”) helps suppress multiple meanings of the same term; in other words, eliminating ambiguity. **NOT** would be represented by the area within the rectangle surrounding both circles, or the “empty” set (“∅”).

Different search engines or tools may provide additional Boolean operators or connectors to create more complex search statements. These may include:

- **Parenthesis:** A Boolean search may include the use of parentheses to force a particular order to the execution of the search, as well as to create more refined and flexible criteria. Any number of logical ANDs (or any number of logical ORs) may be chained together without ambiguity; however, the combination of ANDs and ORs and AND NOTs or BUT NOTs can lead to ambiguous directions. In such cases, parentheses may be used to clarify the order of operations. The operations within the innermost pair of parentheses are performed first, followed by the next pair out, etc., until all operations are completed.
- **Proximity or NEAR/WITHIN operator:** This technique checks the location of terms and only matches those within the specified distance. This is a useful method for establishing relevancy between search criteria, as well as paring down irrelevant matches and obtaining better results (improving precision). Some search engines permit the user to define the order, in addition to the distance of the search terms. For example: budget w/10 deficit would mean “deficit within the 10 words of word budget.” [w/ does not specify order in most systems, only distance; there is another connector that is generally used to specify the particular order. We should ask a vendor if we want to include it here.]
- **Phrase searching:** Some search engines provide an option to search a set of words as an exact phrase, either by typing the phrase in quotation marks (“ ”) or by using a command. When they receive this kind of instruction, the search engine will locate all words that precisely match the search terms, and then discard those which are not next to each other in the correct order. To perform this task efficiently, the index typically will store the position of the word in the document, so the search engine can tell where the words are located.
- **Wildcard operators** (also sometimes referred to as “truncation” or “stemming”). This search capability allows the user to widen the search by searching a word stem or incomplete term. Such a search is typically reflected by a symbol such as a question mark (?), asterisk (*), or exclamation point (!).

The search engine may also allow the user to restrict the truncation to a certain number of letters by adding additional truncation symbols. For example: “Teach??” would find “teaches” and “teacher,” but would not find “teaching.” In addition, some engines will allow for internal truncation such as “wom?n,” which would find “women” or “woman.” The “*” and “!” terms have broader application: for example, hous* would find house, housemate, Houston, household, or other words with the stem “hous.”

B. Probabilistic Search Models

Probability theories are used in IR to make decisions regarding relevant documents. A probabilistic search system is based on a formula that places a value on words, their interrelationships, proximity, and frequency. By computing these values, a relevancy ranking can be determined for each document in a search result. This weighting may be based on a variety of factors:

- Frequency of terms within a document – the more times the term appears, the more weight it carries.
- Location of terms within a document – terms in titles and closer to the top of documents are more heavily weighted.
- Adjacency or proximity – the closer the terms are to each other, the heavier the weighting.
- Explicit or implicit feedback on relevance, in which the top-ranked documents are examined, and used to refine the probabilistic model.¹

Examples of probabilistic search models include Okapi BM25, Bayesian networks, and language models.²

C. “Fuzzy” Search

Boolean and probabilistic search models rely on exact word matches to form the results of a query. Exact matching is very strict: either a word matches or it doesn't. “Fuzzy” search is an attempt to improve recall by matching more than the exact word: fuzzy matching techniques try to reduce words to their core and then match all forms of the word. The method is similar to stemming in Boolean classifiers, discussed above.

Some algorithms for fuzzy matching rely on the understanding that the beginning and end of English words are more likely to change than the center, so they count matching letters and give more weight to words with matching letters in the center than at the edges. Unfortunately, this can sometimes yield results that make little sense (a search for “Tivoli” might bring up “ravioli”).

Many systems allow the user to assign a degree of “fuzziness” based on the percentage of characters that are different. Fuzzy search, or matching, has at least two different variations: finding one or more matching strings of a text, and finding similar strings within a

¹ In explicit relevance feedback, the user codes the top-ranked documents, and only those coded relevant are used to refine the model. In implicit relevance feedback, the top-ranked documents are simply assumed to be relevant.

² See Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze, INTRODUCTION TO INFORMATION RETRIEVAL, 219-252 (2008), available at <http://www-nlp.stanford.edu/IR-book/>.

fixed string set often referred to as a “dictionary.” Fuzzy search has many applications in legal IR including: spellchecking, auto-filling of email addresses, and OCR cleanup.

D. Dimensionality Reduction Systems

Bayesian classifiers are often considered “naïve” because they assume that every word in a document is independent of every other word in the document. In contrast, there is a class of concept learning technologies that rely on the notion that words are often correlated with one another, and that there is value in that correlation. These methods are also referred to as “dimensionality reduction techniques” or “dimension reduction systems.”

These systems recognize there is redundancy among word usage and take advantage of that redundancy to find “simpler” representations of text. For example, a document that mentions “lawsuits” is also likely to mention “lawyers,” “judges,” “attorneys,” etc. These words are not synonyms, but they do share certain meaning characteristics. The presence of any one of these words would be suggestive of a common theme. Documents that mentioned any of these terms would likely be about law. Conversely, in searching for one of these words, one might be interested in finding a document that did not contain that exact word, but did contain one of these related words.

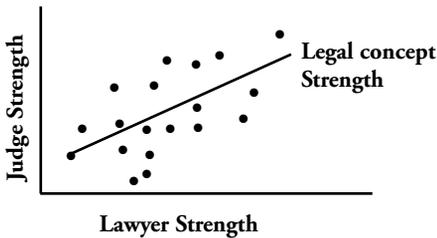


Figure 1. Dimension reduction - the original dimensions of “Lawyer” and “judge” are combined into a single dimension. Each point in the graph represents a document. Its location in the graph shows how much the document is related to each dimension.

The figure above illustrates the kind of relationships identified by such systems. The word “lawyer” tends to occur in the same context as the word “judge.” Each document has a certain strength along the “lawyer” dimension, related, for example, to how many times the word “lawyer” appears. Similarly, documents have strength along the “judge” dimension, related, for example, to how many times the word “judge” appears. These systems find a new dimension that summarizes the relationship between “lawyer” and “judge.” In this example, we are reducing the dimensions from two to one.

Mathematically, we can then describe documents by how much strength they have along this dimension and not concern ourselves with its strength along either the original “lawyer” or “judge” dimensions. The new dimension is a summary of the original dimensions, and the same thing can be done for all words in all documents. We can locate documents along these new, reduced, dimensions or we can represent words along these dimensions in a similar way.

Similarly, multiple words can be represented along dimensions; instead of having just one summary dimension, we can have many of them. Instead of describing a document by how it relates to each of the words it contains, as is done with Vector Space Models,³ we

3 See Roitblat, *supra* note 48.

can describe the document by how it relates to each of these reduced dimensions. Latent Semantic Indexing (“LSI,” also called “Latent Semantic Analysis”) is the most well-known of these dimension-reducing techniques, but there are others, including neural networks and other kinds of statistical language modeling.

These techniques are similar to one another in that they “learn” the representations of the words in the documents from the documents themselves. Their power comes from reducing the dimensionality of the documents. They simplify representation, and make recognizing meaning easier.

For example, a collection of a million documents might contain 70,000 or more unique words. Each document in this collection can be represented as a list of 70,000 numbers, where each number stands for each word (i.e., the frequency with which that word occurs in that document). Using these techniques, one can represent each document by its strength along each of the reduced dimensions.

One can think of these strengths as a “meaning signature,” where similar words will have similar meaning signatures. Documents with similar meanings will have similar meaning signatures. As a result, the system can recognize documents that are related, even if they have different words, because they have similar meaning signatures.

E. Machine Learning Approaches

There are two main types of machine learning: Unsupervised and Supervised. Unsupervised learning is performed using a large set of examples, without any additional human input. In Supervised Learning methods, the learning examples are tagged individually by a user, and the learning process relies heavily on these examples. Both Supervised and Unsupervised learning may use dimension-reduction techniques as described in Section D.

1. Unsupervised Machine Learning (Statistical Clustering)

Systems may use statistics or other unsupervised machine learning tools to recognize the category to which certain information belongs. The simplest of these is the use of “statistical clustering.” Clustering is the process of grouping together documents with similar content. There are a variety of ways to define similarity, but one way is to count the number of words that overlap between each pair of documents. The more words they have in common, the more likely they are to be about the same thing.

Many clustering tools build hierarchical clusters of documents. Some organize the documents into a fixed number of clusters. The quality or “purity” of clustering (i.e., the degree to which the cluster contains only what it should contain) is rarely as high as that obtained using custom built taxonomies or ontologies, but since they require no human intervention to construct, clustering is often an economical and effective first-pass at organizing the documents in a collection. One of the major hurdles when deploying clustering is that there are no objective measures of “clustering quality.” Some systems improve the quality of clusters that are produced by starting with a selected number of clusters, each containing selected related documents. These selected documents then function as “seeds” for the clusters. Other related documents are then joined to them to form clusters that correspond to their designer’s interests. Then, additional documents are added to these clusters if they are sufficiently similar.

2. Supervised Machine Learning

In the context of text categorization, the general objective of machine learning is to generate a classifier that can automatically classify new untagged documents accurately and efficiently based on a small set of tagged exemplar documents. There are many learning algorithms that can be used in the classifier paradigm, including, but not limited to, Naïve Bayesian, Artificial Neural Networks, Support Vector Machines, and Logistic Regression. The choice of which algorithm to use depends on the nature of the task at hand, the type of data, the characteristics of the learning process, etc.

Selecting the set of exemplar documents that will be used to build a classifier is a key challenge of supervised machine learning – it must ensure comprehensive coverage of the population of documents while minimizing the size of this training set to control costs.

One common approach to selecting the training set is the use of active learning. Active learning is a widely researched field, within which several methods and technologies have been developed and tested.⁴

The premise of active learning, and the main differentiator between this approach and other types of machine learning, is that the learning process is conducted in a number of iterations. The algorithm selects a small sample of documents for each iteration. Each sample is tagged by an expert user, and the tags are fed back to the algorithm as a training set. The algorithm learns from these new tags and generates another sample for the next iteration. The process continues until sufficient learning has occurred and the algorithm can accurately predict the user's classification decisions.

The main advantage of the active learning technique is that it enables the algorithm to make an informed decision as to which documents are to be included in the next sample. The basis for this decision becomes more and more informed as the number of iterations increases. The objective of active learning is to optimize learning performance by choosing sample documents that provide the maximum contribution to the training of the classifier. In comparison to random sampling, active learning dramatically reduces the number of documents needed for the training stage.

F. Concept and Categorization Tools: Thesauri, Taxonomies, and Ontologies

To deal with the problem of synonymy, some systems rely on a thesaurus, which lists alternative ways of expressing the same or similar ideas. When a term is used in a query, the system uses a thesaurus to automatically search for all similar terms. The combination of query term and the additional terms identified by the thesaurus can be said to constitute a “concept.”

The quality of the results obtained with a thesaurus depends on the quality of the thesaurus, which, in turn, depends on the effort expended to match the vocabulary and usage of the organization using it. Generic thesauri, which may attempt to represent the

⁴ For a general discussion and comparison of various active learning techniques, see S. B. Kotsiantis, *Supervised Machine Learning: A Review of Classification Techniques*, 31 *INFORMATICA JOURNAL* 249 (2007), available at http://www.informatica.si/PDF/31-3/11_Kotsiantis%20-%20Supervised%20Machine%20Learning%20-%20A%20Review%20of...pdf; G. Cormack and M. Mojdeh, *Machine Learning for Information Retrieval: TREC 2009 Web, Relevance Feedback and Legal Tracks*, 18th Text Retrieval Conference (TREC 2009) Proceedings (2010), <http://trec.nist.gov/pubs/trec18/papers/uwaterloo-cormack.WEB.RF.LEGAL.pdf>.

English language or are specialized for particular industries, are sometimes available to provide a starting point, but each group or organization has its own jargon and own way of talking that require adjustment for effective categorization. In America, for example, the noun “jumper” is a child’s one-piece garment. In Australia, the noun “jumper” is a sweater. In America, a 3.5 inch removable disk device was called a “floppy” during its heyday. But in Australia, it was called a “stiffy.”

Taxonomies and ontologies are also used to provide conceptual categorization. Taxonomy is a hierarchical scheme for representing classes and subclasses of concepts. The figure below shows a part of a taxonomy for legal personnel. Attorneys, lawyers, etc., are all types of legal personnel. The only relations typically included in a taxonomy are hierarchical or inclusion relations. Items lower in the taxonomy are subclasses of items higher in the taxonomy. For example, the NAICS (North American Industry Classification System) is one generally available taxonomy that is used to categorize businesses. In this taxonomy, the category “Information” has subclasses of “Publishing,” “Motion Picture and Sound Recording Industries,” and “Broadcasting.”

One can use this kind of taxonomy to recognize conceptual relationship among these different types of personnel. If a category includes law personnel, then any document that mentions attorney, lawyer, paralegal, etc., should be included in that category. Like thesauri, there are a number of commercially available taxonomies for various industries.

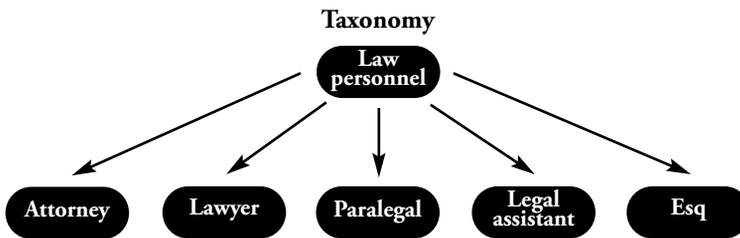


Figure 2. A simple taxonomy for law personnel.

Predefined taxonomies exist for major business functions and specific industries. It may be necessary to adapt these taxonomies to one’s particular organization or matter.

An ontology is a more generic species of taxonomy, often including a wider variety of relationship types than are found in the typical taxonomy. An ontology specifies the relevant set of conceptual categories and how they are related to one another. The figure below shows part of an ontology covering subject matter similar to that described in the preceding taxonomy. For clarity, only a subset of the connections between categories is shown. According to this ontology, if the category includes attorneys, the user may also be interested in documents that use words such as “lawyer,” “paralegal,” or “Esq.” Like taxonomies, ontologies are most useful when they are adapted to the specific information characteristics of the organization.

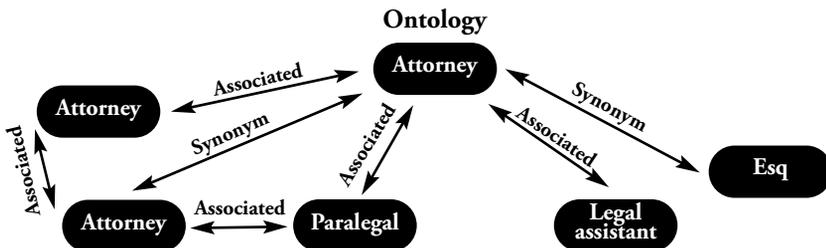


Figure 3. A section of an ontology of legal personnel.

Taxonomies, ontologies, and thesauri are all knowledge structures. They represent explicit knowledge about some subject. An expert writes down the specific relations she knows about. Although there are tools that help the expert create these structures, they still tend to represent only the information the expert can explicitly describe as important.

The structure of the thesaurus, taxonomy, or ontology can be used as the organizing principle for a collection of documents. Rules are derived that specify how documents with specific words in them are related to each of these categories, and the computer can then be used to organize the documents into the corresponding categories.

These rules can be created explicitly, or they can be created using machine learning techniques. Explicit rules are created by knowledge engineers. For example, one rule might include a Boolean statement like this: (acquir* or acquisition or divest* or joint venture or alliance or merg*) and (compet* or content or program*) that specifies the critical words that must appear for a document to be assigned to the “merger” category. The effectiveness of rules like these depends critically on the ability of the knowledge engineers to guess the specific words that document authors actually used. Syntactic rules may also be employed by some systems. For example, a system may only look for specific words when they are part of the noun phrase of a sentence.

G. Presentation/Visualization and Social Networking Tools

Presentation and visualization software technologies may incorporate search and retrieval functionality that may be found to have useful applications. These technologies can organize information (e.g., emails) so that a searcher can more efficiently study the search topic (including finding relevant emails). They also are good at highlighting patterns of “social networks” within an organization that would not necessarily be apparent by more traditional searches. Subject to some exceptions, the results of any search and retrieval query can be presented in a variety of forms, including as a:

1. List – items in sequence, for example messages ordered by sent date
2. Sort – sortable items aggregated into rows by columns, for example messages by sender
3. Group – items categorized or totaled, for example count of messages by sender
4. Cluster – items in groups organized by spatial proximity, for example relevant groups spiraling out to less relevant groups
5. Tree – items in parent/child hierarchy, for example, folder and subfolder(s)
6. Timeline – items arrayed by a time element, for example a list/group of items arrayed by sent date
7. Thread – items grouped by conversation
8. Network – items arrayed by person, for example a diagram of message traffic between sender(s) and recipient(s)

9. Map – items plotted by geography, for example items plotted by city and state of origin
10. Cube – items in a multidimensional pivot table; including, table, group, timeline, and tree functionality

In practice, a searcher can load search results into a presentation technology for an organized view and then drill down to access discrete items of particular interest or concern. This often iterative process may help a searcher to learn more about, act on, and manage search results.

THE SEDONA CONFERENCE COMMENTARY ON ACHIEVING QUALITY IN THE E-DISCOVERY PROCESS*

*A Project of The Sedona Conference Working Group on
Electronic Document Retention & Production (WG1)*

Author:

The Sedona Conference

Editors-in-Chief:

Jason R. Baron

Maura R. Grossman

Jeffrey C. Sharer

2009 Editors-in-Chief:

Jason R. Baron

Macyl A. Burke

Senior Contributing Editor:

Thomas Y. Allman

Thanks go to all who participated in the dialogue that led to this Commentary.

We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.

PREFACE AND ACKNOWLEDGEMENTS

Welcome to another major publication in The Sedona Conference Working Group Series (“WGS”). This is the 2013 Edition of *The Sedona Conference Commentary on Achieving Quality in the E-Discovery Process*, a project of our Working Group on Electronic Document Retention & Production (WG1). The public comment version was first published in May 2009. In the intervening time, the subject of what constitutes a *quality process* in carrying out e-discovery has only grown in recognition and importance. This Commentary has received recognition by having been cited to date both in an influential federal court opinion,[†] as well as in over 20 law reviews and hundreds of legal blogs and websites.

The present Commentary recognizes that the exponentially increasing volume of electronically stored information (ESI) that may be implicated in litigation, investigations, and regulatory activities requires fundamental changes in thinking and practice on the part of the legal profession. As outlined here, these include greater reliance on automated methods to gauge the quality of document productions (including the use of sampling and other forms of measurement), as well as increased attention to project management associated with the e-discovery process. This Commentary is intended to be read in conjunction with *The Sedona Conference Commentary on the Use of Search and Information Retrieval Methods in E-Discovery* (2013), as well as with a forthcoming Sedona Commentary on the subject of defending the e-discovery process.

This 2013 Edition incorporates many of the suggestions and updates provided by a new editorial team formed after the Annual Meeting of WG1 in Austin, Texas in the fall of 2011. I wish to acknowledge the contributions of Jason R. Baron, with whom Maura R. Grossman and Jeffrey C. Sharer took the leading role in revising and updating the prior version, as assisted by Macyl Burke, Todd Elmer, Joe Looby, James Sherer, and Paul McVoy. On behalf of The Sedona Conference, I want to thank the editorial team and all the WG1 members involved in devoting their time and attention during the editing process.

We fully understand that the matter of what constitutes best practices in maintaining quality in a particular legal case will necessarily be subject to change, given the accelerating pace of technological developments with which the law is struggling to keep up. If you wish to submit any further comments, please visit our website at <https://thesedonaconference.org> and join the online discussion forums, or email us at info@sedonaconference.org.

Kenneth J. Withers
Deputy Executive Director
The Sedona Conference
December 2013

[†] *Nat'l Day Laborer Org. Network v. U.S. Immigration Customs and Enforcement*, 877 F. Supp. 2d 87, 109 n.14 (S.D.N.Y. 2012) (Scheidlin, J.).

TABLE OF CONTENTS

Executive Summary.....268

 I. Introduction272

 II. Achieving Quality Through Project Management &
 Better Measurement274

 III. Applying Quality Measures in E-Discovery.....286

Conclusion298

Appendix: Sampling 101 for the E-Discovery Lawyer300

EXECUTIVE SUMMARY

The legal profession has passed a crossroads: When faced with a choice between continuing to conduct discovery as it had “always been practiced” in a paper world – before the advent of computers, the Internet, and the exponential growth of electronically stored information (ESI) – or alternatively embracing new ways of thinking in today’s digital world, practitioners and parties acknowledged a new reality and chose progress. But while the initial steps are completed, cost-conscious clients and overburdened judges are increasingly demanding that parties find new approaches to solve litigation problems. The central aim of the present Commentary is to refine practitioners’ awareness about a variety of processes, tools, techniques, methods, and metrics that fall broadly under the umbrella term “quality measures” and that may be of assistance in handling ESI throughout the various phases of the discovery workflow process. These include greater use of project management, sampling, machine learning, and other means to verify the accuracy and completeness of what constitutes the “output” of e-discovery. Such collective measures, drawn from a wide variety of scientific and management disciplines are intended only as an entry point for further discussion, rather than an all-inclusive checklist or cookie-cutter solution to all e-discovery issues.

While the notion of expressly building in and accounting for “quality” might once have appeared as a somewhat novel idea in the legal discovery context, there is no shortage of competing ideas about quality methods and techniques in the world at large. These include the philosophies and combined works of such individuals as Joseph Juran on quality control,¹ W. Edwards Deming on *total quality management* (TQM) and *statistical process control* (SPC),² Armand V. Feigenbaum on *total quality control* (TQC),³ Phil Crosby on *zero defects*,⁴ Bill Smith of Motorola and *Six Sigma*,⁵ as well as *Capability Maturity Model Integration* (CMMI)⁶ and a host of Japanese *lean* methods, including Kaizen (continuous improvement).⁷

As used in this Commentary, the term *quality control* involves the specific procedures, tools, and techniques that ensure the maintenance of high quality throughout various stages of the e-discovery process (i.e., while people are conducting specific tasks). *Quality assurance*, on the other hand, refers to the methods and metrics used at the end of the process to assess and ensure that an e-discovery process has been performed reasonably and as expected.

Risk of sanctions aside, there are at least four reasons for assessing the quality of an e-discovery process:

1. The failure to employ a quality e-discovery process can result in a failure to uncover or disclose relevant evidence, which can affect the outcome of litigation.

1 JURAN’S QUALITY CONTROL HANDBOOK (J.M. Juran & Frank M. Gryna ed., 4th ed. 1988).

2 W. Edwards Deming, *OUT OF THE CRISIS* (1982).

3 Armand V. Feigenbaum, *TOTAL QUALITY CONTROL* (1961).

4 Philip B. Crosby, *Quality Is Free: The Art Of Making Certain* (1979).

5 See Paul Keller, *SIX SIGMA DEMYSTIFIED* (2005); KPMG, *Six Sigma in the Legal Department: Obtaining Measurable Quality Improvement in Discovery Management* (2006), available at http://www.lean6sigma.vn/index2.php?option=com_docman&task=doc_view&gid=27&Itemid=43 (last accessed Nov 1, 2013); see also “Six Sigma,” https://en.wikipedia.org/w/index.php?title=Six_Sigma&oldid=552846107 (last accessed Nov. 1, 2013).

6 See Mary Beth Chrissis, et al., *CMMI: Guidelines For Process Integration And Product Improvement* (2d. ed. 2006).

7 See M. Imia, *KAIZEN: THE KEY TO JAPAN’S COMPETITIVE SUCCESS* (1986); see also “Continual Improvement Process,” https://en.wikipedia.org/w/index.php?title=Continual_improvement_process&oldid=540550886 (last accessed Nov. 1, 2013); G. Taguchi, *INTRODUCTION TO QUALITY ENGINEERING* (1986); see generally, William Truscott, *SIX SIGMA: CONTINUAL IMPROVEMENT FOR BUSINESSES* (2003) (comparatively summarizing many of the above “quality” methods).

2. An inadequate e-discovery process may result in the inadvertent production or disclosure of privileged and/or confidential information.
3. Procedures that measure the quality of an e-discovery process allow timely course correction and provide greater assurance of accuracy and completeness, especially of innovative processes.
4. A poorly planned or executed effort can cost more in the long run if the deficiencies ultimately result in motion practice or require that e-discovery efforts must be redone.

Thus, the identification and use of best practices in collection, processing, review, and production are essential. Lawyers will always be free to decide how they choose to practice law; however, the courts and clients will decide who wins and loses. Winning in e-discovery increasingly means adopting lean, efficient, and effective business practices that satisfy parties' legal obligations and can withstand judicial scrutiny. These practices include (but are not limited to) using project management and appropriate measures of quality to reduce cost and mitigate risk.

Of course, the reasonableness of a party's discovery process must be evaluated on a case-by-case basis in the context of the proportionality factors set forth in Rule⁸ 26(b)(2)(C),⁹ and no practitioner should assume that any single practice, process, or quality-checking measure is appropriate in any and all circumstances.

The discussion in this Commentary is based on the following guiding principles:

Principle 1. In cases involving ESI of significant scope and complexity, the attorney in charge should utilize project management tools and exercise sufficient leadership to ensure that his or her legal team follows a reasonable process to identify potentially responsive material.

The discovery phase of litigation is best conducted under the active leadership of an attorney, acting individually or as the Team Leader,¹⁰ who is responsible for overseeing the full e-discovery process using project management tools as well as other skills and techniques. The Team Leader (and/or the team he or she leads) should have sufficient experience in the various phases of e-discovery to effectively execute the requisite management duties. Realistically, of course, it may be necessary to delegate responsibility for various phases of the process, but the Team Leader should obtain regular status updates, maintain frequent contact with the team at all levels, and ensure the effective and appropriate communication of information amongst team members.

⁸ For the sake of brevity, the Federal Rules of Civil Procedure will not be shortened to the commonly used abbreviations "Fed. R. Civ. P." or "FRCP" when referenced in the body of the text of the Commentary on Achieving Quality. However, they may occasionally be referred to simply as "the Rules" in a broad or general context. Further, when individual rules are referenced, they will simply be referred to by their numerical indicator preceded by the word "Rule."

⁹ See The Sedona Conference *Commentary on Proportionality in Electronic Discovery* (2013), <https://thesedonaconference.org/download-pub/1778>.

¹⁰ See discussion, *infra* at Part II.A.2.

Principle 2. Parties should employ reasonable forms or measures of quality at appropriate points in the e-discovery process, consistent with the needs of the case and practitioners' legal and ethical responsibilities.

A producing party must make a “reasonably diligent search for emails and other electronic documents”¹¹ sought in discovery, and parties are required to work cooperatively to formulate discovery plans that are embodied in pre-trial discovery orders.¹² These processes can be enhanced by applying reasonable measures of quality, such as: using various forms of sampling at different phases of the process; testing the results of an e-discovery process to determine whether those results are reliable; adopting reconciliation measures for different phases of the e-discovery process; and inspecting results to verify and report whether discrepancies were noted. Qualified individuals should adequately document any quality measures that were used for later support of the e-discovery process.

Rule 26(g)(1), which requires the certification of reasonableness and good faith when requesting and responding to discovery, necessitates a form of quality assurance by counsel based on an appropriate level of attention paid to ensuring accurate results. This is especially true given the exponential increase in the volumes and sources of ESI in the average case, and the new discovery tools and processes that rely on automated means of collecting, filtering, searching, and reviewing massive amounts of potentially responsive data. As discussed below, reasonable project management often entails using various forms or measures of quality at different phases of the process.

Principle 3. A thoughtful and well planned e-discovery “process” should enhance the overall quality of the production in the form of: (a) reducing the time from request to response; (b) reducing cost and burden; and (c) improving the accuracy and completeness of responses to requests.

This Commentary endorses a wide range of quality processes aimed at adding value while reducing cost and effort. A well-designed e-discovery process will employ sound project management practices that tailor the process to the specific case circumstances; use iterative and adaptive procedures and approaches that allow for learning and correction; and, where appropriate, employ statistically sound metrics to monitor course and obtain valid measures of the accuracy and completeness of the e-discovery effort.

Principle 4. Cooperation and greater transparency among parties can be key ingredients to improving quality in e-discovery. Parties should confer early in discovery, including, where appropriate, exchanging information on any quality measures that may be applied.

Generally, cooperation and greater transparency among parties throughout the discovery process can significantly contribute to ensuring quality, maintaining best practices, and reducing claims of spoliation in complex e-discovery. The discovery phase should not be a place for extended argument and advocacy. Rather, discovery should be viewed as an opportunity for cooperation and transparency, in the spirit of Federal Rule of Civil

11 *Wingnut Films, Ltd. v. Katja Motion Pictures Corp.*, No. CV 05–1516–RSWL SHX, 2007 WL 2758571, at *5 (C.D. Cal. Sept. 18, 2007).

12 *Board of Regents of Univ. of Neb. v. BASF Corp.*, No. 4:04CV3356, 2007 WL 3342423, at *5 (D. Neb. Nov. 5, 2007) (emphasizing the duty to cooperatively plan discovery and affirmatively certify reasonableness and good faith as part of an “open discovery process”).

Procedure 1.¹³ The appropriate time for advocacy and argument should arise once discovery is completed – not before or during discovery. See *The Sedona Conference Cooperation Proclamation* (2008) promoting “open and forthright information sharing, dialogue (internal and external), training, and development of practical tools.” It is a fundamental mission of The Sedona Conference to persuade requesting and producing parties that collaboration and dialogue on matters concerning ESI are appropriate, if not necessary.

As they relate to e-discovery, Federal Rule of Civil Procedure 26(f) and similar requirements to “meet-and-confer” are best viewed as a process. The process should start as early as practicable and should extend through the entire discovery lifecycle – identification, preservation, collection, processing, search, review, and production – including discussions, where appropriate, on which search and review processes or technologies will be used and what quality steps will be taken to ensure that these tools have adequately captured responsive documents. The thrust of the 2006 amendments to the Federal Rules of Civil Procedure and a consistent theme that has emerged in the myriad local rules, guidelines, and pilot programs that have been introduced in recent years, is open and forthright sharing of information by all parties during the discovery process, and “removing contentiousness as much as practicable.”¹⁴

13 FED. R. CIV. P. 1 (providing that the rules governing procedures in civil actions “should be construed and administered to secure the just, speedy, and inexpensive determination of every action and proceeding”).

14 *Board of Regents of the University of Nebraska v. BASF Corp.*, No. 4:04CV3356, 2007 WL 3342423, at *5 (D. Neb. Nov. 5, 2007). Examples of pilot programs and special local rules include: Seventh Circuit Electronic Discovery Pilot Program (May 2011), available at http://www.discoverypilot.com/sites/default/files/Principles8_10.pdf (last accessed Nov. 1, 2013); *In re Pilot Project Regarding Case Management Techniques for Complex Civil Cases in the Southern District of New York*, Case No. 1:11-mc-00388-LAP (S.D.N.Y. Nov. 1, 2011) (standing order designating the case for inclusion in the Project), available at http://www.nysd.uscourts.gov/cases/show.php?db=notice_bar&cid=261 (last accessed Nov. 1, 2013); *Default Standard for Discovery, Including Discovery of Electronically Stored Information, U.S. District Court for Delaware* (Dec. 2011), available at <http://www.dcd.uscourts.gov/sites/default/files/Chambers/SLR/Misc/EDiscov.pdf> (last accessed Nov. 1, 2013); and *Guidelines for the Discovery of Electronically Stored Information, U.S. District Court for the Northern District of California*, (Nov. 2012), available at http://www.cand.uscourts.gov/filelibrary/1117/ESI_Guidelines.pdf (last accessed Nov. 1, 2013).

I. INTRODUCTION

“For the rational study of the law the black[letter] man may be the man of the present, but the man of the future is the man of statistics and the master of economics.”

- Oliver Wendell Holmes, *The Path of the Law* (1897)

Oliver Wendell Holmes was right, in ways he could not have imagined more than a hundred years ago. Over a decade into the 21st Century, the legal profession faces exponentially increasing volumes of ESI with all of its attendant complexity.¹⁵ As characterized by one federal District Court judge: “With the rapid and sweeping advent of electronic discovery, the litigation landscape has been radically altered in terms of scope, mechanism, cost, and perplexity. This landscape may be littered with more casualties than successes. ...”¹⁶ Continuing to practice law in the same way – and, particularly, continuing to approach the discovery process in a manner divorced from best practices drawn from other disciplines, without concern for measures of “quality” – increasingly will be a recipe for failure.

Not surprisingly, the case law,¹⁷ industry literature,¹⁸ and law firms seeking a “comparative advantage”¹⁹ have increasingly emphasized the importance of quality measures involved in e-discovery. In *Victor Stanley v. Creative Pipe I*,²⁰ for example, one of the reasons a party failed to carry its burden of excusing inadvertent production of privileged documents was the party’s failure to demonstrate, *inter alia*, “quality-assurance testing.”²¹ Our discussion of the role of quality builds upon Principle 11 of the Sedona Principles:

A responding party may satisfy its good faith obligation to preserve and produce relevant electronically stored information by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information.²²

15 Discovery Search, 17 RICH. J.L. & TECH. 9 (2011).

16 *PSEG Power N.Y., Inc. v. Alberici Constructors, Inc.*, No. 1:05-CV-657 (DNH/RFT), 2007 WL 2687670, at *1 (N.D.N.Y. Sept. 7, 2007).

17 See *William A. Gross Constr. Assoc., Inc. v. American Mfrs. Mutual Ins. Co.*, 256 F.R.D. 134, 134 (S.D.N.Y. 2009) (“This Opinion should serve as a wake-up call to the Bar in this District about the need for careful thought, quality control, testing, and cooperation with opposing counsel in designing search terms or ‘keywords’ to be used to produce emails or other electronically stored information”); *In re Serquel Products Liability Litig.*, 244 F.R.D. 650, 662 (M.D. Fla. 2007) (“Common sense dictates that sampling and other quality assurance techniques must be employed to meet requirements of completeness” [and] [i]f [defendant] took such steps, it has not identified or validated them.”).

18 See, e.g., Ashish Prasad, Kim Leffert, and Shauna Fulbright-Paxton, *Cutting to the Document Review Chase*, BUSINESS LAW TODAY (Nov./Dec. 2008), available at <http://www.abanet.org/buslaw/blt/2008-11-12/prasad.shtml> (last accessed Nov. 1, 2013) (highlighting need for quality control during document review); KPMG, *The Case for Statistical Sampling in E-Discovery* (Jan. 2012), available at <http://www.kpmg.com/us/en/IssuesAndInsights/ArticlesPublications/Documents/case-for-statistical-sampling-e-discovery.pdf> (last accessed Nov. 1, 2013); Deloitte, *eDiscovery Sampling—The What, Why, How, When and Who*, DISCOVERY MANAGEMENT DIGEST – Q3 2012 (2012), available at http://www.deloitte.com/view/en_US/us/Services/Financial-Advisory-Services/Deloitte-Discovery-Financial-Advisory/748a90e606bb9310VgnVCM2000001b56f00aRCRD.htm (last accessed Nov. 1, 2013).

19 Thomas E. Stevens and Wayne C. Matus, *Gaining a Comparative Advantage in the Process*, NAT’L L. J., Aug. 25, 2008

(suggesting use of quality-assurance techniques such as sampling to check results), available at <http://www.pillsburylaw.com/sitefiles/publications/3e2d33870915b82e23983452bae0f7d.pdf> (last accessed Nov. 1, 2013).

20 250 F.R.D. 251 (D. Md. 2008); see also *Mt. Hawley Ins. Co. v. Felman Prod., Inc.*, 271 F.R.D. 125 (S.D. W.Va. 2010).

21 *Victor Stanley I*, 250 F.R.D. at 262. Both *Victor Stanley I* and this Commentary were cited approvingly in *NDLON v. U.S. Immigration Customs & Enforcement Agency*, 877 F. Supp. 2d 87, 109 nn.14 & 15 (S.D.N.Y. 2012).

22 See *The Sedona Principles: Best Practices, Recommendations & Principles for Addressing Electronic Document Production* (2d ed. 2007) <https://thesedonaconference.org/download-pub/81>.

The Commentary to Sedona Principle 11 notes that:

Depending on the nature of the sources of data, both manual and automated procedures for collection may be appropriate in particular situations. Whether manual or automated, the procedures must be directed by legal counsel to assure compliance with discovery obligations. ...

... Regardless of the method chosen, consistency across the production can help ensure that responsive documents have been produced as appropriate.²³

Achieving “consistency across the production” requires sound project planning and oversight, which is often enhanced by integrating measurements of quality within the overall e-discovery process. As management expert Peter Drucker once declared, “If you can’t measure it, you can’t manage it.”²⁴

The ultimate goal in discovery is to identify, collect, and cull documents and ESI from a larger data universe – and to subsequently search for, retrieve, and produce the relevant or responsive, non-privileged materials using tools or methods (whether automated, human, or some combination of the two) that are reasonable and proportional under the circumstances of the case. Where appropriate, the process should incorporate some form of metrics to quantify the accuracy and completeness of the resulting output.

There is no single, “best way” through the e-discovery maze. In determining resource allocation for a given matter, practitioners should weigh the risk of overlooking relevant or privileged material against the advantages of automation, efficiency, and cost savings. In particular, strategic and tactical decisions about how to go about locating relevant evidence in a body of collected ESI (either before or during the review process) are critical.²⁵ This Commentary is not a comprehensive roadmap covering all possible uses of quality measures and metrics throughout the e-discovery process. The creativity of vendors and the bar will ensure that the development and application of quality techniques will continue to advance. Nor is there any bias toward any particular method, tool, or technology, or even a point of view that sampling or other types of quality measures are invariably required in every type of litigation. The solutions to the problems created by scale do not lie exclusively in technology – which is merely a tool – but rather in the effective use of technology by professionals skilled in team leadership, project management, and quality control and assurance.

Following this introduction, Part II.A. discusses the importance of a thoughtful, well-defined e-discovery process, and the need for team leadership and skilled project management. Part II.B. makes the case for why quality matters, and provides examples of five general measures of quality as benchmark guidance. Part II.C. summarizes judicial approaches to sampling. Part III provides selected examples of how quality measures may be applied in various phases of e-discovery, including data collection, review, and production. For the reader so inclined, the Appendix contains a primer on statistical sampling.

23 *Id.* at 58 cmt. 11.c.

24 See David A. Garvin, *Building a Learning Organization*, in HARVARD BUSINESS REVIEW ON KNOWLEDGE MANAGEMENT 46, 70 (HBS Press 1998).

25 To that end, the present Commentary dovetails, and should be read in conjunction with, The Sedona Conference, *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery* (2013), 15 Sedona Conf. J. 217 (2014), which more narrowly focuses on new ways of thinking about issues involving search and information retrieval.

II. ACHIEVING QUALITY THROUGH PROJECT MANAGEMENT & BETTER MEASUREMENT

A. THE IMPORTANCE OF PROJECT MANAGEMENT

1. The Need for a Well-Defined Process

Before embarking on any complex e-discovery project, it is important to recognize, first and foremost, the importance of the process that manages the task, whether it involves “simple” human review or the application of automated tools and more sophisticated techniques. Successfully meeting the challenges posed by large and heterogeneous document collections in e-discovery requires a range of contributions: from people, technology, methodology, and so on. Technologically advanced tools, however “cutting edge” they may be, will only yield successful outcomes when used by people who (a) understand the tools, (b) understand the circumstances and requirements of the case, (c) use thoughtful and well-defined methods, and (d) measure their results for accuracy and completeness. The first step, then, is to develop a thoughtful framework and process within which the applicable methodologies can be applied.

An effective process will usually include most, if not all, of the following key process elements that provide the groundwork for the effective application of technology:

Leadership	Someone who is assigned the responsibility for ensuring that the process reflects a reasonable, good-faith effort to be complete and accurate;
Tailoring	Tailoring of the process to the specific size, risks, needs, and circumstances of the particular case or investigation;
Expertise	Incorporating and drawing upon the appropriate range of expertise required to meet and accomplish the goals set for the particular process, in a timely and cost-effective manner;
Adaptability	Design of an iterative and adaptive process that allows for learning, course correction, and refinement as the project unfolds;
Measurement	Employment of reasonable and appropriate process metrics to monitor progress and ensure consistent and high-quality results;
Documentation	Documentation of the process to permit coordination and communication within the discovery team – and to increase defensibility should the e-discovery effort be subsequently challenged;
Transparency	Explanation of the selection, design, implementation, and measurement of the process in a clear and comprehensive manner to the relevant fact-finder, decision maker, tribunal, or regulator, as well as to opposing counsel, as appropriate;
Cooperation	Solicitation – and incorporation to the extent possible, and within the bounds reasonable effort and advocacy – of input from the requesting party.

2. Project Management and the Need for a Team Leader

The overall quality of any e-discovery effort will be enhanced with increased attention to project management – a discipline popularized by Henry Gantt with roots in a variety of fields, including construction, engineering, and defense.²⁶ In a nutshell, project management “is the discipline of organizing and managing resources (e.g., people) in such a way that the project is completed within defined scope, quality, time, and cost constraints.”²⁷ An almost universal key to the success of any project is the appointment of a project leader, whose responsibility is to:

Lead the team in figuring out what the project is (planning, scheduling, and requirements-gathering), shepherding the project through design and development work (communication, decision making, and mid-game strategy), and driving the project through to completion (leadership, crisis management, and end-game strategy).²⁸

The client must clearly and decisively vest power in an e-discovery project Team Leader, empowering that leader to manage the e-discovery efforts of outside counsel and service providers. A designated outside counsel may serve as the e-discovery project Team Leader but, in some cases, in-house counsel may handle the role. To borrow an analogy from the construction field, the project leader is the “Legal Architect.” In deference to familiar litigation terminology and the combination of expertise typically required for this role, in this Commentary, we refer to this individual as the “lead e-discovery attorney in charge,” or, more concisely, the “Team Leader.”²⁹

The Team Leader, working with the client and any service providers, defines the project’s budget, goals and objectives, and develops a plan for achieving the tasks and activities that need to be performed. The Team Leader understands both the substantive and strategic aspects of the litigation. Where possible, the Team Leader has experience with the various phases of e-discovery and, to the extent feasible, should balance his or her role in developing the facts of the case, interviewing witnesses, and related activities, with leadership of the team’s e-discovery efforts. In large cases, it may be appropriate for the legal team to designate a Team Leader whose primary role on the matter is to manage the e-discovery process. And in certain cases, the Team Leader may also coordinate other (non-electronic) discovery tasks in the case.

Among other things, the Team Leader should ensure that there are regular updates and there is effective dissemination of information to all team members. While daily discussions with each team member may not be practicable, a project may require an oversight and reporting structure that ensures daily communication with at least some team members.

Given the highly specialized nature of some e-discovery tasks, such as information processing, competent assistance – including that of third-party vendors – may be essential.

26 In 1917, Henry Gantt developed the “Gantt Chart,” a tool used for scheduling work (first in factories) using a standard format for the display of scheduling information. Kathy Schwalbe, INFORMATION TECHNOLOGY PROJECT MANAGEMENT, 30 (5th ed. 2007). See generally “Project Management,” http://en.wikipedia.org/wiki/Project_management (last accessed Dec. 13, 2013).

27 See Schwalbe, *supra*, n.26, at 30.

28 Scott Berkun, *The Art Of Project Management* 8 (2005).

29 The use of the singular here and *passim* is not intended to exclude the possibility that in certain legal contexts, one or more “lead” attorneys, with clearly defined duties, may play a substantial role in facilitating the overall e-discovery team effort.

The Team Leader (or other counsel participating in the discovery process) will typically be required by court or agency rules governing production to sign a discovery response or certification, an act that has consequences if the court or agency subsequently challenges the diligence and accuracy of the effort.³⁰ Because of the sheer volume of ESI, we practice in “this age of electronic discovery when attorneys may not physically touch and read every document within the client’s custody and control.”³¹ In order to certify that reasonable, good-faith efforts were made to locate and produce responsive material called for by a document or information request, practitioners must be comfortable and knowledgeable about the process(es) applied in identifying and producing the ESI. Where responsibility and accountability for different phases of the discovery process are divided among multiple persons, a certifying practitioner’s reliance on the team’s efforts must be “reasonable.”³² Under such circumstances, team members must communicate effectively to ensure a seamless, reasonable, and defensible overall process.

A more comprehensive discussion of the advantages of project management is beyond the scope of the present Commentary. However, as is increasingly being recognized,³³ the discipline may yet provide lawyers in leadership roles in the discovery process with a coherent framework for managing large-scale e-discovery matters.

B. THE NEED TO MEASURE QUALITY IN E-DISCOVERY

The concept of quality – long an important consideration for many business and manufacturing processes – increasingly is being applied to the e-discovery process. In e-discovery, an assurance of quality focuses on the usefulness of a given task’s results, as measured by the likelihood that a particular tool or method has adequately identified responsive documents and ESI.

What are the barriers to successful adoption of quality measures in e-discovery? For one thing, the variety and changing needs of cases often lead to an *ad hoc* approach to discovery management.³⁴ Moreover, in many cases, important roles are either delegated directly to third parties, or rely upon complex processes and software managed by third parties; however, either situation requires appropriate quality tests to measure those processes.³⁵ Finally, there is no universally accepted standard for a “quality” outcome. In

30 Under Federal Rule of Civil Procedure 26(g)(1)(B), for example, an attorney (or unrepresented party) that signs a discovery response is deemed to have certified that, to the best of the person’s knowledge, information and belief “formed after a reasonable inquiry” that the response is not interposed for an improper purpose and neither unreasonable nor unduly burdensome. See *Mancia v. Mayflower*, 253 F.R.D. 354, 360 (D. Md. 2008) (noting that Rule 26(g) requires “approaching the process properly ... in accordance with the letter and spirit of the discovery rules”); accord *Cache La Poudre Feeds v. Land O’Lakes*, 244 F.R.D. 614, 628-630 (D. Colo. 2007) (noting that counsel is required to undertake a reasonable investigation to identify and preserve relevant materials in course of responding to discovery requests).

31 *Qualcomm, Inc. v. Broadcom Corp.*, No. 05cv1958-B (BLM), 2008 WL 66932, at *9 (S.D. Cal. Jan. 7, 2008), *vacated in part and remanded*, No. 05CV1958-RMB (BLM), 2008 WL 638108 (S.D. Cal. Mar. 5, 2008).

32 Where in-house counsel is the one to perform the efforts to search and produce documents and ESI, certifying counsel may rely “on assertions by the client and on communications with other counsel as long as that reliance is appropriate under the circumstance.” Advisory Committee Note, FED. R. CIV. P. 26, Subdivision (g) (1983). Cf. *Qualcomm, supra*, n.31. One way that certifying counsel can do this is to forward the draft discovery request, response, or objection to in-house counsel (and where appropriate, the relevant in-house IT representative), to facilitate review and confirmation that the submission is complete and accurate before it is served or filed.

33 See, e.g., Association of Corporate Counsel, ACC VALUE CHALLENGE: GUIDE TO PROJECT MANAGEMENT (2011). <http://www.acc.com/advocacy/valuechallenge/toolkit/loader.cfm?csModule=security/getfile&pageid=1293786&page=/index.cfm&qstring=&title=ACC%20Legal%20Project%20Management%20Primer> (last accessed Nov. 1, 2013); Jeane Thomas & Ben Hawksworth, *Lessons Learned, Master Mining: Three Views on EDD Project Management*, LAW FIRM INC. (Mar./Apr. 2006), reprinted in http://www.crowell.com/documents/DOCSOCFKTYPE_ARTICLES_396.pdf (last accessed Nov. 1, 2013).

34 Litigation projects involving custom specifications defined by the user, while highly configurable and flexible, lead to a non-standardized approach that does not lend itself to repeatability and inherent engineered quality.

35 See n.17, *supra*.

traditional discovery, that standard has always been one of “reasonableness,” rather than “perfection”; this Commentary does not argue to the contrary.³⁶

1. Why Does Quality Matter?

Faced with uncertainty, some practitioners question whether assuring quality in e-discovery is really that important. Sanctions aside, we believe that quality is extremely important for at least four principal reasons:

First, failure to employ a quality e-discovery process can result in failure to uncover or disclose key evidence. This reason is the most compelling and, potentially, most important for the parties to a case. A simple example: If search terms are used without quality testing, a party may not find exculpatory or “hot” documents (and their many near duplicates) crucial for convincing an adversary to settle the matter or, for that matter, inculpatory documents that could make settlement a wise strategy for the producing party. Depending on the size and scope of the case, the implications can be expensive and far-reaching.³⁷

Second, a poorly conceived or managed e-discovery process may result in the inadvertent production of privileged or confidential information.³⁸ This common concern often prompts many outside counsel to undertake resource intensive, manual review of electronic documents, with its attendant high cost. Moreover, some aspects of collection and review frequently involve corporate intellectual property (IP), trade secret, and otherwise confidential ESI. A quality process will identify these items and designate them for protective treatment earlier rather than later in the process.

Third, e-discovery processes that incorporate measures of quality are more defensible because they provide metrics – and, if properly implemented, allow for course correction and refinement. A Team Leader who measures discovery process quality as discovery progresses, or reasonably soon thereafter, is well positioned to make any necessary modifications. If there are mistakes, or if a systemic or systematic error is discovered, an informed Team Leader may be able to modify the process in midstream before there is a challenge to the production.

Fourth, poor quality can lead to deficiencies in production, ensuing motion practice, and e-discovery efforts that must be redone. Each step leads to increased costs. While historically, higher quality meant higher cost; that is not necessarily the case today. It is often less expensive to engineer quality into the process than to add it after the fact. In fact, many quality programs have begun incorporating measurement, tracking, and savings reports into the process as a measure of success.

³⁶ While the reasonableness standard is not a bright-line rule and lends itself to ambiguity, confusion and, at times, disputes; it also is a standard that is easily adaptable to the numerous different discovery contexts and cases. *See, e.g., Pension Committee of the University of Montreal Pension Plan v. Banc of America Securities*, 685 F. Supp. 2d 456 (S.D.N.Y. 2010), *overruled in part on other grounds by Chin v. Port Authority of New York & New Jersey*, 685 F.3d 135 (2nd Cir. 2012); *see also* FED. R. CIV. P. 26(g)(1) (certification of discovery responses made “to the best of the person’s knowledge, information, and belief *formed after a reasonable inquiry*”) (emphasis added); *Rinkus Consulting Group, Inc. v. Cammarata*, 688 F. Supp. 2d 598, 613 (S.D. Tex. 2010) (“Whether preservation or discovery conduct is acceptable in a case depends on what is *reasonable*, and that in turn depends on whether what was done – or not done – was *proportional* to that case and consistent with clearly established applicable standards. As Judge Schiendlin pointed out in *Pension Committee*, that analysis depends heavily on the facts and circumstances of each case and cannot be reduced to a generalized checklist of what is acceptable or unacceptable.”) (emphasis in original); Craig B. Shaffer, “*Defensible*” *By What Standard?*, 13 SEDONA CONF. J. 217, 222 (2012) (“While a defensible e-discovery plan is not held to a standard of perfection, [FRCP 34] does require a party to undertake reasonable efforts to identify and produce responsive, non-privileged material in its possession, custody or control.”).

³⁷ *See, e.g., Qualcomm, supra*, n.31.

³⁸ *See Victor Stanley I, supra*, n.21.

These points raise a related issue that parties involved in e-discovery should understand: the distinction between “quality control” and “quality assurance.”

Quality control involves engineering quality into a process. It uses procedural safeguards built into a process to ensure high quality throughout, and it focuses on the execution of specific tasks.

For example, if someone is moving 10 items from point A to point B, the quality-control step in that process would be to count the items before the move, and then again after the move, to ensure that all of the items were transferred (i.e., a “reconciliation”). This Commentary suggests that quality control should be built into the e-discovery process.

Quality assurance, on the other hand, typically takes place after a process is complete and involves an assertion as to what was done, how well it was done, and whether the output met a certain predetermined standard. Quality assurance generally refers to the procedures designed to serve as the basis for certification and reliance. Because quality assurance often involves an intensive, third-party audit of process and activities, it can be much more intensive and expensive – particularly in a large, complex e-discovery project. For many e-discovery projects, a full-scale quality-assurance process may not be practical, financially feasible, or proportionate in the circumstances of the case, and thus, would be unwarranted. In such cases, it may be appropriate to consider a more targeted approach. Five principal measures of quality are especially useful in regard to e-discovery.

2. Five Measures of Quality

a. Judgmental Sampling

Accountants typically use a form of judgmental sampling as a form of quality control to find material misstatements where there are many similar financial transactions; this sampling method can greatly reduce the cost of an audit while maintaining its integrity. A typical example is accounts payable, where a large number of invoices pass through and are subjected to the same procedures, thus allowing representative samples to be drawn.³⁹

What can be described as “judgmental sampling” also has been used in traditional e-discovery.⁴⁰ Attorneys often select a few folders of electronic documents coded by a particular reviewer to determine whether the reviewer is making the correct responsiveness calls. After reviewing the judgmental sample, the more senior attorney may, based on the exercise of informed judgment, request additional samples or require a heightened second-level review if the perceived error rate is unacceptable. A judgmental sample, unlike a statistically valid sample, does not permit one to make assertions about the entire population from which the sample was drawn with statistical confidence, but can nevertheless be very helpful.⁴¹

39 Janet L. Colbert, *Audit Sampling*, INTERNAL AUDITOR (Feb. 1, 2001), at 27.

40 The distinction between “judgmental” and “statistical” sampling is discussed in more detail in the Appendix (“Sampling 101 for the E-Discovery Lawyer”).

41 The reviewer may be aware of, and take into account, the source of the documents, the size of the population, the types of information at issue, and the degree to which the results conform to other sources of oral, written, or physical testimony or evidence.

This method is especially useful as a quality check on discretionary processes, such as collection and review. For example, an experienced professional can review a report from the collection phase that lists evidence sources collected for each custodian, and may identify gaps in the collection (e.g., a hard drive was not collected for a particular custodian). The professional selects this and other apparent exceptions and has staff research the anomaly and perform remedial action (e.g., collect the hard drive), or annotate the collection report (e.g., note that the individual does not use a personal computer), as appropriate.⁴²

The selection of keywords as search terms for responding to discovery requests is a special form of judgmental sampling that is based on many factors, including prior knowledge as well as educated guesses with respect to what a collection of ESI may contain.⁴³ There will always be some measure of informed judgment involved with the selection of search or filtering criteria at various phases of e-discovery.

Notably, the initial results produced by human judgmental sampling can be – and, increasingly, are expected to be⁴⁴ – further refined and improved through greater use of iterative processes.⁴⁵

b. Independent Testing

Third-party professionals can be retained to examine a process or approach and report on whether results can be replicated and confirmed. One such example might be automated or highly technical processes, such as data processing, searching, or computer-assisted coding, that have been challenged on the basis of reliability and accuracy. Thus, a native file processing application (i.e., software that converts files from their “native” or “proprietary” form into a generic form, such as *.tiff images for further processing) might be tested to validate (or invalidate) the software’s reported efficacy at extracting files from an email container, accurately displaying such files for review, and indexing the searchable text in such files.

Currently, these “black box” technologies often are described only by what they can do – not by what they cannot do – leaving the industry in a *caveat emptor* situation. This is expected to change, as there have been a few widely reported deficiencies in the capability of certain native file processing technologies to completely render email, extract embedded objects, search compound documents or containers, extract metadata, and the like.⁴⁶

42 Similarly, an experienced professional can review a report from the review phase that lists the documents marked or reviewed, by reviewer, per hour, or relevant documents per custodian, and from a high level, may be able to identify outliers or other unusual patterns.

43 See *Sedona Search Commentary*, *supra* n.25, *passim*.

44 See, e.g., *Natl. Day Laborer Org. Network v. Immigration and Customs Enforcement Agency*, 877 F. Supp. 2d 87 at 110 (noting “research showing that, in many contexts, use of keywords without testing and refinement (or more sophisticated techniques) will in fact not be reasonably calculated to uncover all responsive material”); *In re Sequel Products Liability Litig.*, 244 F.R.D. 650, 662 (M.D. Fla. 2007) (“[W]hile key word searching is a recognized method to winnow relevant documents from large repositories, use of this technique must be a cooperative and informed process. ... Common sense dictates that sampling and other quality assurance techniques must be employed to meet requirements of completeness.”).

45 See *Victor Stanley I*, 250 F.R.D. at 262 (“Selection of the appropriate search and information retrieval technique requires careful advance planning by persons qualified to design effective search methodology. The implementation of the methodology selected should be tested for quality assurance; and the party selecting the methodology must be prepared to explain the rationale for the method chosen to the court, demonstrate that it is appropriate for the task, and show that it was properly implemented. In this regard, compliance with The Sedona Conference *Best Practices for Use of Search and Information Retrieval* will go a long way towards convincing the court that the method chosen was reasonable and reliable.”). See also George L. Paul & Jason R. Baron, *Information Inflation: Can The Legal System Adapt?*, 13 RICH. J.L. & TECH. 10 at 50 (2007) (suggesting iterative protocols be used).

46 One additional method to verify systems and processes is to employ “known sample” testing, where systems and processes are applied to a known collection, with defined characteristics, to measure the results. This is useful in establishing the basic scope and functionality of search systems and review protocols. In particular, this method is used in by legal-service providers to test prospective reviewers against a “test folder” of already-coded documents, to establish how well the reviewers can absorb and apply a given review protocol.

Similarly, in some circumstances, it may be appropriate to call upon third-party professionals to recommend a test design that would allow independent verification that a production has met an agreed-upon standard of accuracy. Ideally, the general framework for such independent testing would be agreed upon by the parties at the outset of the discovery phase of the litigation or investigation (see below, on the question of when quality measures and metrics may be appropriate).

c. Reconciliation Techniques

Reconciliation to account for the impact of a process (i.e., comparing inputs to outputs) has long been used in the fields of accounting, manufacturing, and engineering and may have applicability to the e-discovery process. E-discovery process reconciliation might involve comparing what volume of email or ESI enters a process, what remains in a process (after, for example, deduplication), and what exits a process. This could help determine whether email or other files were handled correctly, or identify gaps in the process that may have resulted in the omission or incomplete handling of files.⁴⁷

d. Inspection to Verify and Report Discrepancies

Inspection and observation of participants in the e-discovery process resemble the original form of quality control, which was part of the apprenticeship model for training junior attorneys. Inspection is especially useful as a quality check on processes such as collection and review. The deployment of seasoned experts to inspect and observe the performance of tasks by less-experienced staff can improve quality on a project. For example, during collection, it is often advisable to have senior legal and technical participants involved in the initial custodian interviews. Experienced staff can coach less-experienced staff, as well as further define or refine the process (such as by improving the custodian interview questionnaire). During the review phase, it is equally advisable to improve quality by having a seasoned reviewer observe the review and provide guidance. This should be done on a frequent and iterative basis, as reasonable and appropriate.

e. Statistical Sampling

The concept of measuring the quality of the output of a process by sampling is not new. Acceptance sampling,⁴⁸ for example, was used by the U.S. military to test the quality of bullets manufactured for use during WWII, to spot design defects, and ultimately, to improve production.

Today, some form of acceptance sampling is used as a quality-control tool by virtually every large and medium-sized manufacturing company in the world. For example, Boeing's instructions to its suppliers require that they "perform either 100% inspection or acceptance sampling for receiving inspection."⁴⁹ This can be seen as a fundamental type of quality control, which is composed of many methods, depending upon the application.

⁴⁷ Reconciliations from the phases before and after native file processing can be extremely complex.

⁴⁸ First introduced in the 1920s by Walter A. Shewhart, this approach is referred to as "Statistical Process Control," and uses statistical tools to observe the performance of the production line to predict significant deviations that may result in rejected products. By maximizing the efficiency of war production, William Edwards Deming popularized the use of this quality control method.

⁴⁹ The Boeing Co., *Quality Assurance Manual*, at D8-4890.105 (2001), available at <http://www.boeing.com/terms/d8-4890105.pdf> (last accessed Nov. 1, 2013).

Standards organizations, such as the International Standards Organization (ISO), a network of the national standards institutes of 157 countries, have created procedures and guidelines for acceptance sampling. Statistical sampling also is required by many government agencies to test the quality of a given population of products.⁵⁰

As noted, statistical sampling permits statements to be made about the population from which the sample was drawn with statistical confidence and is helpful when one wants to get as close to the truth as possible, but time and cost prohibit the testing of each item, or such testing is technically infeasible. It is a “scalable solution,” one that works well regardless of the size of the sampled population.

In the e-discovery context, statistical sampling can serve as a check on the effectiveness of search and review efforts at identifying responsive information and correctly coding documents, whether the efforts are manual or technology-assisted. A party could identify a random sample of documents that the review method did not identify as potentially responsive and review them for responsiveness. By doing so, the party can obtain an estimate of the number of responsive documents remaining in the set of documents that were not selected for further review. Based on the results of such testing, the producing party can take informed actions to improve its review process to close the gap between what was identified as responsive and what was actually responsive.⁵¹

Statistical sampling also can be used to measure the probable error rate for a project, a key custodian’s documents, or even for a specific document reviewer. An acceptable error rate can be defined and document groups with error rates above this threshold can be re-reviewed and retested until the results meet or exceed the quality standard.⁵²

Whether or not a given error rate is truly indicative of an effective review effort will depend upon the prevalence of responsive material in the document collection to begin with. Gauging that will generally require taking into consideration the number of documents in the collection that the search or review method deemed to be responsive, and the proportion of those so deemed that are actually responsive (i.e., the precision). Only when measured in relation to these data points (i.e., prevalence and precision) will an error rate, in the set of documents deemed non-responsive, have any meaning regarding the effectiveness of the review effort (i.e., serve as a gauge of whether or not the review effort has, in fact, succeeded in identifying a reasonably high proportion of the responsive material in the document collection). Put another way, in deciding on what is and is not an acceptable error rate, one should begin by deciding what is and is not an acceptable level of recall and then translate that level of recall into the corresponding rate of error in the set of documents deemed non-responsive.

The size of the sample (and associated review time and cost) required to ascertain that a review effort has met a given standard of quality and, more specifically, met a given

50 For example, the U.S. Department of Agriculture uses sampling plans to test the quality of most of the products it regulates and has codified sampling requirements in Title 7, § 43 of the U.S. Code. The Internal Revenue Service allows the use of sampling estimates by taxpayers to determine amounts where other estimates are not feasible.

51 See *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012), 2012 WL 1446534, at *2 (S.D.N.Y. Apr. 26, 2012) (protocol utilizing iterative sampling); *In re Actos (Pioglitazone) Products*, No. 6–11–md–2299, 2012 WL 3899669 (W.D. La. July 30, 2012) (same).

52 For example, a review team has a set of documents to review for a particular custodian, say 100,000. A smaller team can then review a random sample of those documents to determine how many were incorrectly coded (either as responsive or non-responsive). The team would also decide an acceptable error rate, e.g., 5%. If the number of incorrectly coded documents for the particular custodian is above this threshold, the team can review additional documents, conduct searches if there was a common error in coding (e.g., an issue missed by a number of reviewers). The team then would then perform another random sample to determine whether the error rate is now within the established, acceptable threshold.

level of recall, will vary with the values taken by a number of parameters, some of which are discretionary – e.g., the standard of quality to be met and the confidence level sought in establishing that the standard has been met – and some of which are not (e.g., the prevalence of responsive material in the document population). Some parameter settings will entail very large sample sizes; others will entail relatively small samples. Parties should take into account the specific circumstances of a given review effort in arriving at a sampling design that will strike the optimal balance between the information provided and the effort and resources required to obtain that information.

3. When are Quality Measures and Metrics Appropriate?

While some or all of the preceding techniques can be helpful in measuring and improving the quality of an e-discovery process, there is no universal consensus on when and how they should be applied – or even what constitutes a quality process. Nor can one expect to simply transfer “off-the-shelf” industrial techniques to e-discovery.⁵³ That is why, to the extent it is feasible, parties may want to discuss how quality will be assured and, if appropriate, measured.

Because it is not practical to apply every (or perhaps any) metric to every step in every case, legal teams must prioritize and determine which quality measures should be applied, and when, based on various factors that include, among others, the value of the claims or damages, the size and complexity of the case, and the time and resources the parties have to expend on implementing quality-control measures. Clients, opposing parties, courts, and regulators may have an important say in the selection of these procedures as well.

In the case of complex matters – for instance, an SEC or criminal investigation of a high-ranking corporate officer – a team may opt or be required to perform judgmental sampling, independent tests, and inspections as every phase unfolds; and then do statistical sampling, reconciliations, and independent testing at the end of the process, because cost considerations in such cases will often be secondary to ensuring that the process is as accurate and complete as it can be.

At the other end of the spectrum may be a very small case where the use of quality metrics likely would be light, if they were used at all. In other words, the marginal utility of an assessment of quality must be weighed against, and should be proportional to, the burdens and costs involved, and the anticipated benefits.

Some other questions to consider in assessing the appropriateness of the use of quality methods and metrics are:

- Whether and how quality measurement tools and methods will be used? Will quality be measured during the culling process or further downstream, such as during relevancy or privilege review, or production, or at multiple stages of the project?

⁵³ In analyzing the quality of a given review process in identifying responsive documents, one may need to factor in a scale of relevance – from technically relevant, to material, to “smoking gun” – in ways which have no direct analogy to the industrial-based processes referenced above. Indeed, most quality applications assume one is looking at independent and fairly homogenous events (i.e., events of similar or like character), with an implied inference that probability affects each item equally. This may or may not be the case in the e-discovery context, and therefore, it may be important to understand notions of variance and how much variance should be tolerated coming out of a given process.

- Who will be applying the tools and methods? Will they be applied by experts or by non-experts (e.g., attorneys, paralegals, litigation support staff, and/or third parties)? If by non-experts, consideration must be given to the transfer of case-specific knowledge to the users, to the training of the users in effective use of the tools and methods, and to the overall usability of the tools and methods themselves.
- How will the output be used? Will the tools and methods be used for testing or assessment purposes only (e.g., in order to learn more about the target collection, in order to navigate more efficiently through the documents, in order to prioritize the review, and so on), or will they be used to discern what should receive further review and what should not?
- On what part of the collection will the quality measures be used? Will quality only be assessed regarding documents that are potentially responsive, or also upon the documents that are presumptively non-responsive after culling techniques have been applied?
- To what degree is it anticipated that an expert will be required to defend the process, or that another party's expert will attack the scientific validity of the tools and methods?

4. The Need for Documenting a “Quality” Process

Employing the above measures of quality in e-discovery may be prudent, even necessary, in many settings. It is equally important carefully to document the use of such processes; indeed, failure adequately to document the steps taken to sample, test, inspect, reconcile, or verify one's results can affect the litigant's ability to defend its process to the opposing party or to the court. In *United States v. O'Keefe*,⁵⁴ consistent with earlier case law,⁵⁵ the Court noted the importance of providing the Court with appropriate explanations in cases where the parties have not agreed upon the use of keywords and the search task has been performed unilaterally.⁵⁶ In *Victor Stanley I*, the Court found that defendants were “regrettably vague” in describing their approach to keywords, regarding “how they were developed, how the search was conducted; and what quality controls were employed to assess their reliability and accuracy.”⁵⁷ Accordingly, the Court went on to hold that attorney-client privilege had been waived. Challenges in that type of situation should be addressed by persons competent to do so, not merely by conclusory statements of counsel.⁵⁸ Documentation of the process(es) to be employed also may be helpful in situations where courts proactively encourage parties to cooperate in discussing sampling and other protocols to be used as part of overall e-discovery plans.⁵⁹

⁵⁴ 537 F. Supp. 2d 14 (D.D.C. 2008).

⁵⁵ See *Judicial Watch, Inc. v. U.S. Dep't of Justice*, 185 F. Supp. 2d 54, 64 (D.D.C. 2002) (expressing inability to determine if a search was “reasonably calculated” to recover documents where the party failed to explain keywords used).

⁵⁶ *Accord Equity Analytics, L.L.C. v. Lundin*, 248 F.R.D. 331 (D.D.C. 2008) (involving file extensions).

⁵⁷ 250 F.R.D. at 256. See also *Walter A. Gross, supra*, n.15 (highlighting need for attention to be paid to quality control in developing search terms); *Rhoads Indus., Inc. v. Bldg. Materials of Am.*, 254 F.R.D. 216, 224 (E.D. Pa. 2008) (referencing need for “proper quality assurance testing,” citing *Victor Stanley I*).

⁵⁸ 250 F.R.D. at 261 n.10 (to be accomplished by “affidavits or other equivalent information from persons with the requisite qualifications and experience”).

⁵⁹ See, e.g., *Columbia Pictures Indus. v. Bunnell*, No. CV 06-1093FMCJCX, 2007 WL 2080419, at *14, n.32 (C.D. Cal. May 29, 2007) (court “encourages the parties to meet and confer regarding sampling”); *Zurich American Ins. Co. v. Ace American Reins. Co.*, No. 05 Civ. 9170 RMB JCF, 2006 WL 3771090, at *2 (S.D.N.Y. Dec. 22, 2006) (court orders parties to devise a protocol for sampling).

At the outset of any e-discovery process, the Team Leader (or his or her delegate) should determine the documentation standards and controls appropriate for the particular matter, and then re-evaluate those standards and controls as the matter progresses, to ensure their ultimate defensibility. The Team Leader should act under the assumption that every aspect of the process employed could be challenged and, as appropriate, include quality measures designed to answer those challenges in the overall project plan. This entails creating and updating documentation in real time, as decisions are made, to best assure that declarations and other statements regarding the outcome of the process will be adequately supported if and when the need arises at a later date.

C. JUDICIAL APPROACHES TO SAMPLING

The Federal Rules of Civil Procedure contemplate the use of “sampling” as a means of reducing the enormous burdens posed by the vast volumes of ESI in litigation today. For example, in connection with Rule 26(b)(2)(B), the Advisory Committee noted the affirmative role that sampling may play when assessing whether “good cause” has been shown to order production of information from sources identified as not reasonably accessible, stating that “the parties may need some focused discovery, which may include sampling of the sources, to learn more about what burdens and costs are involved in accessing the information, what the information consists of, and how valuable it is for the litigation in light of information that can be obtained by exhausting other opportunities for discovery.”⁶⁰

In the litigation context, the term “sampling” can have a number of different meanings, depending upon how the sample is selected and the purpose for which it is used. Thus, courts distinguish between “judgmental sampling” and “statistical sampling,” the latter of which has been traditionally used primarily in determining “adjudicative facts.”⁶¹

As early as 1963, survey conclusions based on “random sampling”⁶² and the application of probabilistic principles were deemed admissible in court as evidence to establish facts in dispute.⁶³ Thus, for example, statistical sampling has routinely been used in discrimination cases to assess whether discrimination has occurred.⁶⁴ These cases draw on probability theory to determine whether the observed variations at issue may have resulted from chance, or whether they demonstrate a pattern of intentional misconduct.⁶⁵

More recently, in the “light’ cigarettes” class action litigation,⁶⁶ Judge Weinstein concluded that “[s]ampling and survey techniques are well-accepted alternatives for the trial judge facing crippling discovery and evidentiary costs,” and that “[g]reater reliance on

60 In a different context – acknowledging the power of one party to compel a sample from another under certain circumstances not relevant here – Rule 34(a) was amended to state that “[a]ny party may serve on any other party a request to produce and permit the party making the request, or someone acting on the requestor’s behalf, to inspect, copy, test, or sample any designated documents or electronically stored information” in the respondent’s possession, custody or control, and also includes provision for entering on to the premises of the responding party “for the purpose of inspection and measuring ... testing, or sampling. ...” (emphasis added).

61 *Rosado v. Wyman*, 322 F. Supp. 1173, 1180 (E.D.N.Y. 1970), *aff’d*, 437 F.2d 619 (2d Cir. 1970), *aff’d*, 402 U.S. 991 (1971).

62 See Appendix, *infra*.

63 In *Zippo Mfg. Co. v. Rogers Imports, Inc.*, the plaintiff manufacturer of cigarette lighters employed three surveys of the relevant smoking population to demonstrate a likelihood of confusion among consumers in support of its claim for trademark infringement. 216 F. Supp. 670 (S.D.N.Y. 1963). For an interesting and comprehensive discussion of court decisions involving statistical sampling, see Laurens Walker & Joan Monahan, *Sampling Evidence at the Crossroads*, 80 S. CAL. L. REV. 969 (2007).

64 See, e.g., *Castaneda v. Partida*, 430 U.S. 482, 495-96 (1977); *Stewart v. General Motors Corp.*, 542 F.2d 445, 449 (7th Cir. 1976); see also *Capaci v. Katz & Besthoff, Inc.*, 711 F.2d 647, 653-57 (5th Cir. 1983).

65 See *Ageloff v. Delta Airlines, Inc.*, 860 F.2d 379, 383-84 (11th Cir. 1988); see also *G.M. Brod. & Co., Inc. v. U.S. Home Corp.*, 759 F.2d 1526, 1538-40 (11th Cir. 1985) (expert testimony established profit projections based on industry standards).

66 *Schwab v. Philip Morris USA, Inc.*, 449 F. Supp. 2d 992, 1245, 1247 (E.D.N.Y. 2006), *rev’d* on other grounds sub nom. *McLaughlin v. Am. Tobacco Co.*, 522 F.3d 215 (2d Cir. 2008).

statistical methods is required by the profound evolution in our economic communication and data compilation and retrieval systems in recent decades.”⁶⁷ The Court noted that the Supreme Court had recently permitted sampling as a method for the Internal Revenue Service to assess unreported tips by restaurant employees.⁶⁸

As described below,⁶⁹ statistical sampling can help provide assurance of the accuracy of automated or other tools used to reduce the size of a given population of ESI for purposes of review.

Statistical or probabilistic sampling stands in contrast to the less formal “judgmental sampling,” that is often used to facilitate the exercise of discretion by a court or by a party seeking to assess the quality of a process. This is the context referred to in the Committee Note to Rule 26(b)(2)(B) quoted above. For example, courts routinely utilize samples of arbitrary size drawn from a population of potential sources to help them exercise judgment as to the extent to which a party should restore backup tapes, and who should bear the costs of doing so. In the cases of *McPeck v. Ashcroft*⁷⁰ and *Zubulake v. UBS Warburg LLC*,⁷¹ for example, the Courts utilized informal sampling techniques in connection with assessment of the marginal utility of investing resources in the restoration and recovery of ESI from backup tapes. Similarly, in *Quinby v. WestLB AG*,⁷² the Court refused to require a producing party to absorb the full cost of restoring and searching backup tapes where a review of a sample showed that “only a small percentage of the emails produced are relevant.”⁷³

The use of keyword searches to cull or filter relevant information from massive amounts of ESI is another example of the use of judgmental sampling, the effectiveness of which is subject to evaluation by testing. “For example, a producing party could apply a certain set of keywords and/or concepts to cull down a sample of the collection and then analyze the results.”⁷⁴ Trial or pilot runs of combinations of words may be tested in an iterative fashion to extrapolate the effectiveness of the chosen set. In the case of *Clearone Communications v. Chiang*,⁷⁵ for example, the Court noted that an initial effort to modify conjunctive search terms was, in effect, a first step in a “sampling” process to avoid over inclusiveness, or what are commonly referred to as “false positives.” As noted by another Court, “[c]ommon sense dictates that sampling and other quality assurance techniques must be employed to meet requirements of completeness.”⁷⁶ There is a growing recognition that, in most cases, this will require that the producing party review not only those documents that “hit” on selected search terms, but also review samples drawn from those documents that do not hit on such terms in order to identify the rate of “false negatives” and then, in an iterative fashion, to supplement or refine the search terms as necessary to reduce that rate to a level that is acceptable under the circumstances of the particular case.⁷⁷

67 *Schwab*, 449 F. Supp. 2d at 1244-45.

68 *See United States v. Fior D'Italia, Inc.*, 536 U.S. 238, 247-48 (2002); *see also Schwab*, 449 F. Supp. 2d at 1244.

69 *See generally* discussion *infra*, Part III.B.2.

70 202 F.R.D. 31, 34-35 (D.D.C. 2001).

71 217 F.R.D. 309, 324 (S.D.N.Y. 2003).

72 245 F.R.D. 94 (S.D.N.Y. 2006).

73 The court found that the number of relevant documents was “quite low when compared to the volume of documents produced.” *Id.* at 109.

74 *See* Mia Mazza, Emmalena K. Quesada, Ashley L. Sternberg, *In Pursuit of FRCP 1: Creative Approaches to Cutting and Shifting the Costs of Discovery of Electronically Stored Information*, 13 RICH. J. L. & TECH. 11, at 38 (2007) (when results show the use of those combinations did not remove a large volume of relevant information, it validates and supports their use to cull down the remainder of the collection).

75 Civil No. 2:07cv00037TC, 2007 WL 3275300 (D. Utah Nov. 5, 2007), *modified* No. 2:07 CV 37 TC, No. 06-30378, 06-30379, 2008 WL 920336 (D. Utah Apr. 1, 2008).

76 *Seroquel Products Liability Litig.*, 244 F.R.D. at 662; *see also In re Viox Products Liability Litig.*, No. 06-30378, 06-30379, 2006 WL 1726675, at *2 (5th Cir. May 26, 2006) (appellate court urging parties to “adher[e] to a statistically sound protocol for sampling documents” as an aid in reducing the trial court’s burden in reviewing the assertion of privilege on 30,000 documents), *on remand*, 501 F. Supp. 2d 789 (E.D. La. 2007).

77 For example, the *Da Silva Moore* and *Actos* cases involving technology-assisted review, *see* n.51, *supra*, are examples of where the parties have agreed to conduct sampling along these lines.

III. APPLYING QUALITY MEASURES IN E-DISCOVERY

E-discovery involves locating and delivering non-privileged documents and ESI responsive to non-objectionable discovery requests, using a reasonable search method tailored to the needs of the case. The question for the producing party is how best to capture and properly produce this deliverable, what resources should be allocated to the task, and how those resources should be used.

Each case or matter – whether a discrimination claim, patent dispute, or antitrust “second request” – has its own dynamics that will shape the project plan. A series of predictable decisions will be required and each phase of discovery has quality aspects that need to be considered. A non-comprehensive list of examples drawn from select phases of the discovery process follows.

A. THE DATA COLLECTION AND CULLING PHASES

1. Building on Traditional Approaches to Document Collection and Culling

In the days of paper, lawyers knew how to ask for and collect “documents.” Key custodians would be asked to gather their hard-copy documents and files into boxes, which were made available to lawyers or paralegals to review. Practitioners essentially reviewed each and every page for relevance and privilege. This time-worn process admittedly grew more complex in large litigation matters (e.g., antitrust actions or products liability class actions), where tens or even hundreds of thousands of boxes of documents were collected from a corporate enterprise for review by legions of junior associates and contract attorneys.⁷⁸ Much the same process continues to be employed for the review of large bodies of evidence that exist only in hard-copy form.

With the advent of computers, the Internet, network servers, email, and the explosion in types and volumes of ESI, the collection process has had to adapt to the rapid changes and volume considerations involved. Yet, there is still a need to understand what to ask for that potentially could be relevant, what the sources of those items might be, and what key players would best know about the relevant materials. What has changed materially is the need to engage IT and business professionals who are knowledgeable about the sources and locations of ESI within the enterprise, as well as any outsourced storage that may exist in the possession of third parties or in the “cloud.” These professionals will be informed as to what ESI is online, near-line, and off-line; what may be zipped or encrypted; what may be found on backup sources such as CDs, DVDs, virtual storage devices, servers, and removable storage devices (e.g., flash drives, mobile devices, etc.); and in archives of all kinds. These are all potential sources for collection in any e-discovery process that seeks to identify information relevant to a particular litigation or investigation.

2. Applying Measures of Quality to Data Collection and Culling

While it is generally understood, particularly where ESI is involved, that perfection is not the standard (nor even, in most cases, attainable),⁷⁹ parties do have an

78 See, e.g. *Transam. Computer v. IBM*, 573 F.2d 646, 648 (9th Cir. 1978) (refusing to find waiver of privilege where “unique circumstances” existed, requiring production within a three-month period of approximately 17 million pages, creating “monumental” problems because the pages were not “grouped or batched together”).

79 See *Pension Committee*, *supra* n.36.

obligation to conduct e-discovery reasonably, proportionately, and ethically, demonstrating appropriate professional care and judgment. Given the *ad hoc* nature of the process, it is not surprising that objective benchmarks, standards, and regulations specific to the governing of this process do not yet exist. Absent such standards, how does a careful practitioner ensure that a collection workflow is going to be successful?

Quality-control processes employed prior to the review of ESI are an essential element of the “reasonableness” of a party’s discovery efforts; they also support a quality chain-of-custody process for purposes of tracking and documentation. Parties using a well-designed discovery methodology should be able to account, if necessary, for all of the electronic information they collect (as well as identify any potentially relevant ESI they did not collect), even though they may review and ultimately produce only a small fraction of that information. As a general proposition, these quality-control procedures have two main purposes: data accountability and anomaly detection. The complexity associated with multi-location collection from large numbers of custodians, live systems, archives, and forensic images, makes effective quality-control processes essential. Without them, parties are more vulnerable to challenges related to the omission of potentially relevant data, spoliation, conversion of data, or other issues arising in the later review and production phases of e-discovery.

The collection and culling (i.e., initial processing) of ESI can be measured and managed through the gathering and reporting of key metrics. This analysis should be applied as early in the workflow as possible, helping to communicate the details about the composition of the collected ESI. Simple metrics, such as how much and what types of data have been collected for each source, whether custodial or noncustodial, can be very helpful to the planning process.⁸⁰ This can help with early detection of potential issues such as metadata loss, encryption, corruption, unsupported or unknown file types, non-searchable ESI, and other unpredictable issues. It can also help prevent unexpected cost and burden at the time of review and production; for example, by avoiding the selection of a review tool that is unable to handle the specific languages or file types contained in the ESI.

3. Data Collection and Culling: Best Practice Guidelines

The selection, organization, and filtering of ESI through the use of objective criteria (such as date filters and file types) and, in most cases, a search protocol,⁸¹ are critical elements in reducing the volume of information to be reviewed, and thus, the time and cost of the e-discovery process. In addition, keyword search techniques are well known and may be used for this purpose with proper testing. More advanced technologies employ complex algorithms for ESI filtering and organization and may, in some cases, be useful at the collection and culling stages. Regardless of the technology chosen, all filtering methods require a well-defined process. Without these basic steps, the use of any filtering technology will likely result in gross over- or-under inclusion of responsive ESI. The process includes several steps:

- Understanding the composition of source ESI;
- Defining clear goals for the filtering process;

80 For example, if a party has collected an average of 10 GB of ESI from each of several custodians, a custodian with only 1 GB of ESI collected may stand out if he or she was expected to have the same amount. Creating a corporate- or case-level data map in the early assessment stage of a case gives counsel a framework to analyze and make comparisons of the collection metrics.

81 See *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 372 (S.D.N.Y. 2006) (discussing importance of a “search protocol” in assuring a “diligent search” involving a “reasonably comprehensive search strategy”).

- Applying the filter; and
- Testing the outcome.

Understanding the composition of source ESI is critical when filtering the information. There are many possible types of documents in any given ESI collection: some documents may be handwritten; others may be written in one or more languages other than English; some may be compressed (e.g., *.zip files), encrypted, or otherwise protected; and still others may be composed of images without searchable text (e.g., electronic faxes or scanned paper documents). This may result in having to format, convert, translate, or otherwise specially process documents for subsequent filtering or review. For example, scanned paper documents must first be processed through optical character recognition (OCR) software to create searchable text. Without a well-defined process for all file types, some files may be ignored or missed during the filtering phase of the e-discovery workflow. Finally, the handling of handwritten documents should be separately addressed.

Defining clear goals for the filtering process will help achieve the intent of the filtering. First, the team needs to clearly articulate the filtering's intent – such as reduction of volume by exclusion, inclusion, organization, or classification of ESI – so that the appropriate tool can be utilized and the process can be explained.

The application and testing of any filtering process should be iterative and often may need to be repeated until the desired goals are met. It generally is not sufficient to run a filtering tool and trust that it is achieving the desired results without performing any follow-up measurement. Rather, a practitioner must evaluate the outcome of the search, looking for errors in how the filter rules were established or applied. A practitioner may use metrics, such as the number of included or excluded documents by keyword or filtering criteria, to evaluate the outcome. Examining keywords that return high and low numbers of “hits” can uncover issues with how the search was constructed, the choice of terms, or even issues with the data. For example, finding zero search “hits” on a key term or concept may point out that a search term is spelled incorrectly or was not stemmed appropriately, or that many of the documents do not contain searchable text. On the opposite end of the spectrum, finding a term that “hits” on a high percentage of the ESI may indicate that the term is too broad, or may need to be combined with additional qualifying search terms.⁸²

Practitioners also should maintain data accountability through a chain-of-custody process. An initial, important step maintains the identification of the original source of data at each stage of processing, memorializing the file location, the directory and drive mappings on hard drives, and the contents and file counts for each unique source for each custodian. Best practices would include clear documentation of both what was done and what was not done. If Internet email is not processed for review, that should be documented. If a party uses forensic tools to recover deleted emails from a custodian's mailbox, that use should be documented (including the specific tool and the results). By applying a custodian-based view to the data, the party can report the total number of items from all sources applicable to each custodian (e.g., the live email server, email archive, hard drives, network shares, and removable media). Searching the review database for all items

82 Sampling can also be useful in testing the effectiveness of filters, i.e., samples can be drawn and reviewed from both the set of documents hit by a filter and from the set excluded by a filter, so as to provide evidence as to whether the filter casts too wide or too narrow a net. Sample-based testing of filters, in which all data (filtered and unfiltered) is in-scope for testing, allows for informed, iterative development, ideally leading to more effective filters.

associated with a custodian should yield a result that matches the total number of items from the processing report.

A cornerstone of data accountability is derived by establishing the counts of files on media before processing begins.⁸³ From this defined starting point, the party should make adjustments to file counts reflecting processing results for each source:

- Elimination of system files (e.g., based on the National Software Reference Library or “NSRL” filter);⁸⁴
- Deductions for certain file types not processed (e.g., databases);
- Deductions for items that could not be processed (e.g., corrupt or virus-infected files, or documents created by proprietary software);
- Deductions for duplicates not processed; and
- Deductions for items not selected by filters.

In addition, it is also important to note files processed but not indexed, such as encrypted files or images.

The raw data regarding the electronic information is also useful for identifying anomalies in the data collected. Investigating and resolving these anomalies and exceptions in the data can serve as an additional quality-control check that may discover errors or omissions in the collection process. At a minimum, investigating anomalies may help to answer questions about the collection process that other parties or the court may have. In addition, undocumented but indexed exceptions raise a particular concern, as those items are otherwise loaded into the database and appear to be available for searching, analysis, and review.⁸⁵

Email conversion may implicate additional issues. Conversion of email from one format to another is often necessary during the collection and processing of ESI because many service providers do not natively support all email formats. Given the potential for data loss or alteration (for example, loss of formatting, metadata, etc.) arising as a result of such “conversions,” it is prudent to establish a process that will lead to a reliable and defensible result.⁸⁶

Ultimately, of course, the quality and completeness of the collection and culling phases will be directly related to the care and planning – as well as the ability to adapt to changing needs – that are built into the protocol and executed under the leadership of the Team Leader.

83 Note that file counts can, however, be misleading. Container files (e.g., PSTs) can and should be “exploded” for the purpose of fully indexing and understanding the overall data set and data profile.

84 See NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, <http://www.nslr.nist.gov/> (last accessed Nov. 1, 2013).

85 If they are not indexed, any keyword searches performed against the data will not include those items. Therefore a review that is based, at least in part, on the results of keyword searches may overlook these items.

86 Among the questions that could be asked of a service provider with regard to email conversion are: (i) what email formats require conversion for processing; (ii) what conversion software will be used, and for what purpose was the software developed; (iii) what known limitations in the software exist, including the above-described data loss or alteration; (iv) how has the software been tested or vetted prior to selection for use, and how has the conversion process been tested or validated; (v) what is done to investigate or remediate errors; (vi) what errors are typical; (vii) does the conversion process alter date and time values; (viii) does the conversion process preserve email addresses, resolved names, entire email headers, blind copies (bcc’s), etc.; (ix) does the process retrieve email from all containers in the mail store; (x) does the conversion process handle duplicates, email threads, encrypted messages, corrupt messages, foreign languages, Rich Text (RTF) formatted emails, HTML formatted emails; and (xi) does the conversion process handle non-email objects (e.g., calendar entries, contacts, notes, etc.).

B. THE REVIEW AND PRODUCTION PHASES

1. Introduction

Producing parties review documents or ESI for relevance and responsiveness before they are produced, as well as to determine if any privilege or other exemption is applicable.⁸⁷

Effective management of the review phase requires organization and advance planning. Thoughtful structuring of the review process and coding protocols, sufficient attention to staffing, and open communication and collaboration among attorneys, clients, and legal-service providers are hallmarks of well-managed document reviews. Attention to quality control throughout the process is paramount to providing meaningful and cost-effective results.

Traditional large-scale document reviews typically have involved large numbers of individuals of varying expertise and responsibility. Leadership by individuals who will manage the diverse roles and tasks in coordination with the Team Leader (or his or her delegate) is crucial in the following respects:

- Providing clear guidelines governing “knowledge transfer” from the trial team and managers to the document review team and those assisting on a particular review project. Such guidelines also ensure an adequate “knowledge transfer” from those familiar with the documents and data, back to the trial team and senior managers;
- Developing the applicable responsiveness and privilege criteria;
- Determining search methodologies;
- Ensuring consistency, to the extent possible, especially in areas relying on human review and judgment; and
- Providing for objective quality benchmarks to measure the performance and effectiveness of the review process.

2. Using Automated Methods to Reduce the Initial Burden of Review

As discussed in connection with the collection process, and as set forth in the *Sedona Search Commentary*, there are a number of automated tools that can greatly assist in reducing the overall data universe, in at least four fundamental ways: first, by helping to identify only the portion of the universe of collected ESI that is potentially responsive, based on whatever automated search methodologies and protocols are employed; second, by de-duplicating (or tagging) identical or “near” duplicate ESI so that a particular ESI object is reviewed only once; third, by eliminating certain types of files that likely are not relevant (e.g., video and program files); and fourth, by identifying and eliminating obvious spam from the review population.⁸⁸

⁸⁷ Throughout this section, the term “privilege” is used to refer to both the attorney-client privilege and the attorney work product doctrine.

⁸⁸ See *Sedona Search Commentary*, *supra* n.25, at 245.

A number of legal-service providers now offer various forms of automated tools that can significantly reduce the number of electronic documents to be reviewed manually, by extracting the documents most likely to be responsive to a discovery request and leaving the remainder unselected and unreviewed.⁸⁹ Given the rising costs of complying with e-discovery requests, tools that enable a party reasonably, appropriately, and substantially to reduce the amount of ESI that must be reviewed by humans should be embraced by all interested parties – plaintiffs, defendants, the courts, and government agencies.

For example, to conduct an automated search process utilizing statistical sampling, the first step effectively shares and transfers knowledge among counsel, the project management team, and those with knowledge of the corpus of ESI that is the subject of e-discovery. The knowledge gained in this process is used to develop one or more search strategies (e.g., Boolean searches, concept searches, metadata filters, language-based approaches using taxonomies and ontologies, statistical “clustering” techniques, relevance ranking, machine learning, or other proprietary strategies). Once the responsive data set has been characterized, a random sample of categorized material is chosen; reviewers analyze this sample. This random sample will contain both responsive and non-responsive material, and reviewers classify these documents as they normally would in a manual review. The reviewers’ classifications are then compared to the results reached by the chosen categorization method(s). When there is a difference between the determination made by the human reviewer and the categorization method, the legal team reviews the document and decides which is correct. Adjustments are then made to the search strategy. Sometimes the differences require modifications so that a particular type of document is filtered in the future.

This iterative process of sampling and refinement is typically continued until the rate of difference between the automated methods and human review meets the acceptable threshold of accuracy defined for the project. Once it reaches that threshold, the final categorization is run on the entire data set, and the responsive documents can then be prepared for production (subject to any further privilege or other manual review deemed necessary).

Another approach employs a “blended review” process, combining the use of manual review by humans with automated software-based review. Specifically, automated tools can help reduce the workload when manual review is required. For example, where potentially privileged documents are involved, a “privilege” vocabulary identifying attorney or law firm names can be used in conjunction with a vocabulary identifying topically-relevant information. In this way, the likely privileged documents can be identified initially, and then subcategorized by relevancy, allowing the manual reviewers to start with, and focus on, the most pertinent privileged documents.

Once the ESI universe has been reduced to the population of documents that will be reviewed, there are review tools on the market that can help review teams organize and more effectively work with what still may be a considerable volume of electronic data. For example, a number of service providers offer categorization and “clustering” tools that group documents relating to the same or similar topics so they can be reviewed together – accelerating the speed of review and ensuring greater consistency in the treatment of the documents.⁹⁰ Other providers offer email threading, which likewise pulls together related

89 The magnitude of the cost reduction typically depends on the percent of responsive documents in a data population, the number of different issues relevant to the data request or subpoena, and the extent to which human review is applied.

90 See Ramana Venkata & Michael A. Geibelson, *Overcoming E-Discovery Challenges with New Technologies*, 30 L.A. LAWYER, June 2007, at 46, 46 (“Quality control measures, document organization, and batch tagging can significantly lower the cost of e-discovery and help satisfy client demands for finite, predictable e-discovery costs.”), <http://www.lacba.org/Files/LAL/Vol30No4/2388.pdf> (last accessed Nov. 1, 2013).

emails from a given email chain so they can be reviewed together – as opposed to being randomly dispersed throughout the document set and reviewed independently by different reviewers. Still other providers offer software that can rank a document collection from those documents most likely to be relevant to those least likely to be relevant, thereby allowing the legal team to focus its resources and make strategic decisions about what documents require expensive attorney review. Because many of these tools work by extrapolating coding decisions made on a sample set of documents to make predictions about the entire corpus, it generally will be important to have strong quality controls in place during the sample review.

Although all of these review tools are quite useful in reducing the time spent in reviewing ESI, in most present-day litigation there will still, to a greater or lesser extent, be a need to review manually some portion of the document population to determine responsiveness to a particular e-discovery request. Moreover, it is still common for all or some portion of the data to be reviewed a second time (i.e., by a different reviewer) for privilege, or as a quality check on the first-level review, or simply to understand the facts.

To the extent that automated search and retrieval methods are used for reducing the ESI data set to a more manageable size for purposes of review, a party may be called upon to demonstrate to opposing parties, courts, and government agencies that its chosen method and tool accurately captured a reasonable amount of the relevant, non-privileged ESI, and that the prevalence of relevant documents in the remaining, unreviewed and unproduced ESI is acceptably low. *See also* discussion, *supra*, Part II.B.3.

In citing to the *Sedona Search Commentary*, as well as to the federal government's TREC Legal Track research initiative,⁹¹ the Court in *Victor Stanley I* stated that "there is room for optimism that as search and information retrieval methodologies are studied and tested, this will result in identifying those that are most effective and least expensive to employ for a variety of ESI discovery tasks."⁹²

In addition, while the case law in this area will have to evolve to address challenges to particular e-discovery processes, parties employing these methods may support their use by citing past studies and by employing accepted statistical sampling techniques. As noted in an article summarizing legal principles related to e-discovery:

As with any technology, it is imperative to perform frequent, thorough checks to make sure that the searches are working, perhaps by using a sampling method. ... The needs of the litigation at issue should dictate what technology gets used, and how, in order to strike the optimal balance possible between recall and precision. The key to defensibility is that litigants employ these search strategies as part of a reasonable, good-faith, well-documented discovery protocol. Lawyers must understand where the search technology fits into that protocol and have confidence that they have taken measures to ensure the quality of their searches. (footnotes and internal quotes omitted)⁹³

91 *See* <http://trec-legal.umiacs.umd.edu> (last accessed Nov, 1, 2013); *see also Sedona Search Commentary, supra* n.25.

92 250 F.R.D. at 261 n.10. *See also* technology-assisted review cases cited n.51, *supra*.

93 M. Mazza, E. Quesada, A. Sternberg, *supra* n.74, at 33 (litigants looking for a "holy grail" in automated technology will not find it).

Thus, as noted earlier, sampling techniques can be used to establish, at a certain confidence level, that unreviewed and unproduced material in the form of ESI is likely to be non-responsive to a particular discovery request.⁹⁴

3. “Clawbacks,” Rule 502, and Reliance on Automated Methods

In addition to the tools and techniques set forth in Part II, growing volumes of ESI in litigation will require new, creative, and strategic approaches to efficiently manage e-discovery. Together, the “clawback” provision of amended Federal Rule of Civil Procedure 26(b)(5) (allowing for the return or “clawback” of inadvertently produced privileged materials) and Fed. R. Evid. 502 (generally immunizing parties from third-party challenges if their “clawback” agreement is included in a court order) can provide powerful protection against privilege waiver resulting from inadvertent production of privileged material in discovery. In light of these developments, and without employing labor-intensive manual review (except for spot checking for responsiveness and privilege), in appropriate circumstances, counsel may wish to assess their client’s interest in producing potentially responsive documents gathered as the result of automated search and filtering methods.⁹⁵ That is, even with a “clawback” agreement in place, a party may wish to perform sufficient sampling and other quality-control methods to reach a level of comfort on matters of relevance and privilege, but once that comfort exists, rely primarily on the automated methods to determine which documents are produced. This allows the majority of documents to be produced without exhaustive manual review, and traditional manual review will only be employed on documents initially identified through automated means as candidates to be withheld on grounds of privilege. This approach substantially reduces costs and burdens for both producing and requesting parties, and ideally, the pros and cons of such an approach should be discussed at the outset of the discovery process. Adoption of this approach does not, however, obviate the need for employment of measures of quality at every stage.

4. Quality-Control Guidelines for Responsiveness and Privilege

The use of quality-control tools throughout the review process can provide ongoing performance metrics and resolve potential ambiguity in training and instructions. Indeed, when quality checking is combined with training, a resource-intensive review process can be made much more efficient by improving the quality of initial reviews with iterative feedback.

In traditional document review, there are two basic approaches to quality checking first-level reviewers’ coding decisions when dealing with ESI. One approach is to have a second level of review performed – by a senior lawyer or using trained computer software – on some or all of the coded ESI. Another is to run statistical analyses of coded documents to check for consistency across reviewers, and then conduct a targeted, second-level review only where there are unexplained statistical variations.

⁹⁴ The same techniques can be used to determine, at a certain confidence level, that the produced documents are, in fact, relevant.

⁹⁵ See generally Advisory Committee Note to FED. R. EVID. 502: “Depending upon the circumstances, a party that uses advanced analytical software applications and linguistic tools in screening for privilege and work product may be found to have taken ‘reasonable steps’ to prevent inadvertent disclosure”; see also *Rhoads Indus., supra*, n.57 (citing to Advisory Committee Note).

However, as recent research⁹⁶ and common sense confirm, members of any legal team can disagree on relevance assessments of specific ESI. The standard for first-level reviewers (often contract attorneys) should therefore never be perfection, especially given that “reasonable minds” can differ on exactly what is or is not relevant to a particular request. To further enhance the quality of first-level review, guidance should continually be refined to assist the reviewers in getting at least the “easy” documents correct, and in coming as close as possible to the desired result on “close-call” determinations.

Review for privilege can require an even more nuanced legal analysis and, as such, can be more expensive per document than review for relevance or confidentiality. Complexities include, for example, the use of email chains and internal legal department communications.⁹⁷ Incorporating processes that safely minimize the number of records being reviewed for privilege will lead not only to a more efficient, cost-effective review, but also to faster turnaround for production and higher quality privilege logs. These processes might include:

- Creating a “potentially privileged” set of documents that obtains more scrutiny from more experienced reviewers. This can involve the use of file extensions, document, sources, keyword searches, metadata filters, and any internal designations of privilege. (Note, however, that all footers can contain the words “confidential” or “privileged.”)
- Identifying structured data sources containing data that is never in the hands of an attorney and is not created for any litigation purposes. However, companies and counsel need to be aware that such data sources may contain information that is nonetheless prohibited from disclosure by operation of law, regulation, contract, or that is confidential or proprietary for other reasons.⁹⁸

Quality control is a constant process throughout the production phase of litigation. After the review of an initial subset of ESI has been completed and all reviewers are deemed to be performing adequately, the team should implement a quality-control protocol to apply on a going-forward basis. This protocol should be adjusted to accommodate the reviewers’ growing understanding and new developments in the case.

For example, the system should address the percentage of ESI checked, as well as the methods for selecting ESI to be checked. It is not always necessary for a team to review 100% of the responsive ESI population a second time; instead, sampling measures may be appropriate. To guard against inadvertent production of privileged ESI, in some reviews, a complete check is made of ESI with attorney names (and/or other terms commonly

96 Research shows substantial variance in how human subject volunteers “assess” relevance or non-relevance in response to a hypothetical production requests under FED. R. CIV. P. 34. See generally Maura R. Grossman and Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, 17 RICH. J.L. & TECH. 11 (2011); Maura R. Grossman and Gordon V. Cormack, *Inconsistent Responsiveness Determination in Document Review: Difference of Opinion or Human Error?*, 32 PACE L. REV. 267 (2012); Jason R. Baron, David D. Lewis, & Douglas W. Oard, *TREC 2006 Legal Track Overview*, Fifteenth Text Retrieval Conference (TREC 2006) Proceedings, available at <http://trec.nist.gov/pubs/trec15/papers/LEGAL06.OVERVIEW.pdf> (last accessed Nov. 1, 2013).

97 See, e.g., *Muro v. Target Corp.*, 243 F.R.D. 301 (N.D. Ill. 2007) (intermingling of non-privileged communications with privileged communications in an email chain); *In re Vioxx Prods. Liab. Litig.*, 501 F. Supp. 2d at 815 (adopting Special Master Report involving review of in-house lawyer communications); *Victor Stanley I*, 250 F.R.D. at 262 (failure to carry burden of excusing inadvertent production of privileged documents because of lack of demonstration of, *inter alia*, “quality-assurance testing”).

98 Where parties or their counsel are unwilling or unable to determine that a group of records are completely barren of privileged material, they may wish to consider an offer to produce these documents under a “quick peek” agreement.

associated with privileged documents). Prudence also suggests close checking of the ESI in the production queue that originates from the files of custodians known to work closely with attorneys.

Consistency checking may assist practitioners in evaluating the review. Every discrete object in the review population should ultimately be coded as responsive or not. Accordingly, a check can be done to locate any ESI that exists in the database without the requisite coding. Rather than waiting until the end of the review, this check can be done on a rolling basis as the review proceeds (e.g., at the conclusion of review for each relevant custodian). In addition, consistency in coding across like documents is important; and practitioners should adopt measures to ensure, or at least increase the likelihood, that duplicate or near-duplicate documents are being coded in the same fashion.

Practitioners also should identify other inconsistent combinations, and regularly search the database for such errors, both to correct them and to determine their source so as to correct course and avoid creating more errors. Listed below are a few examples of consistency searches, with the caveat that they will necessarily vary from review to review:

- ESI coded as responsive without coding for potential privilege (where known subsets of the population are considered privileged);
- ESI coded as potentially privileged without coding for privilege type;
- ESI coded as non-responsive and noteworthy;
- ESI coded as non-responsive and potentially privileged;
- ESI coded as in need of redaction, but no redaction has been applied.⁹⁹

Additional common practices for quality control during the review process include:

- Comparing coding among reviewers on a common subset of ESI and providing individual or group feedback following the results;
- Developing a system for early review and assessment before reviewers get too far into review (e.g., the first 100 coded ESI objects are automatically sent to the second-level review team for assessment);
- Selecting a group of documents by running a search by reviewer, specifically targeting codes that will give a broad idea of how each reviewer is handling certain codes;
- Analyzing the daily tracking sheet (i.e., data log) and/or service provider software-generated statistics to identify the frequency of second-level changes to first-level reviewers' coding decisions, and targeting areas of frequent disagreement;
- Conducting early assessment of ESI marked "privileged" to identify any misunderstandings about privilege, and if necessary, providing feedback;

⁹⁹ State-of-the-art review platforms offer support for structures – such as mutually exclusive tag groups – that help enforce these consistency rules during the review process. As they evolve, they may be support for many more such structures (“if it gets this tag, it must also have this one”) that make it easier for inconsistencies to be identified. When available, these should be understood and fully utilized in the design of the review.

- Conducting detailed quality review of “potentially noteworthy” ESI;
- Conducting detailed quality review of key custodians or high-level custodians (e.g., the CEO) to assess the need for early client feedback;
- Running tests to identify logically inconsistent coding (e.g., a “responsive” tag but no responsive coding category or issue tag is selected);
- Tracking rates of review (i.e., the time taken per reviewer per document), and investigating any outliers compared to peer reviewers;
- Gathering metrics on the overall review process itself, including ESI or documents reviewed per hour, “pages” per hour, sampling for congruence, and total costs.

The hosting service provider and/or litigation support project manager can assist with statistical tracking of review efforts (both for substance and pace) for all levels of review. Service providers offer different options for review tracking and can provide reports that track, among other things, the number of “pages” reviewed, the number of ESI objects coded with a particular tag, the hourly average of “pages” reviewed, and the number of times second-level reviewers change the coding of a particular first-level reviewer.

At least three additional types of quality issues may arise in connection with document review:

- Review of data “exceptions.” It is not unusual for different providers and software applications to have difficulty processing obscure, password-protected, encrypted, or corrupted records in the data set. Some data may also be unreadable, for example, characters may be scrambled or random. Service providers typically refer to such documents as “exceptions.” Service providers should be able to provide “exception reports” regularly throughout the review and address these issues on an ongoing basis.

Because exceptions can take considerable time to resolve, waiting until the end of the review to address these files can be a mistake because it can interfere with timely production. If the provider is not able to remedy the problem, third-party consultants may be called in to process the data with different tools. Some password-protected materials can be “cracked,” and others may need to be returned to the client for input. Finally, the content of the exception data should be considered. If it can be determined from the file name or type that it is not likely to contain relevant information, a party should consider whether to inform the opposing party or government agency that the file will not be processed further.

- ESI in foreign languages. Another category of ESI that would commonly be set-aside during review is foreign-language materials. It may be necessary to obtain translations (informal translations using Internet tools may suffice for the purposes of preliminary

review), so time should be built into the production schedule to get these materials translated, reviewed, coded, and processed for production, as required.

- Loose electronic media. If certain types of media have not been made part of the primary review database (such as video or audio tapes, disks, DVDs, CDs, etc.), it is important to build time into the review schedule for review of those materials prior to the final production.

5. Preproduction Quality Checking

Practitioners should employ both preproduction cleanup measures and additional quality-assurance checks, as warranted. Once the review of a set of documents and ESI is complete, the service provider and the team should perform preproduction tasks and conduct final quality checks to ensure that the documents are properly queued, numbered, and labeled for transfer to the final production medium. Although many teams implement quality-control measures throughout the review process, it is important to develop a plan for a final quality review of the preproduction set, which can include performing an appropriate analysis to verify that all ESI coded as responsive and not privileged has been queued for inclusion in the production. This review might also include checking for:

- Inadvertent production of privileged ESI, to ensure that no documents marked “privileged” have been loaded onto the production media. The team may also want to run relevant attorney names or “potentially privileged” terms through the preproduction set as a final check. It is important that the team verify that redactions are properly applied in all production formats.
- Non-responsive ESI, to ensure that no non-responsive data has been loaded onto the production media.
- Inconsistent or illogical coding of ESI, to identify all documents containing inconsistent or illogical coding (e.g., the reviewer checked a responsive issue tag and also checked “non-responsive”).
- Data formats and labeling, to confirm that all data on the production media has been loaded in the required or agreed-upon format.¹⁰⁰ Care should also be taken to ensure that the documents on a disk or other production media are properly labeled, including any necessary confidentiality or FOIA designations, and that the disks or other production media themselves are also properly labeled for production.

6. Final Quality Checking at Production

While quality checking should be undertaken at each stage of the review process, a final check is appropriate prior to production:

- Checking to confirm that the number of files being produced matches the expected number of files (i.e., file count and

¹⁰⁰ The service provider should also check the production media to ensure that all data will be readable and that no data has become corrupted.

reconciliation), to protect against the inadvertent inclusion of extra (possibly hidden) files.

- Each production to opposing counsel should be quality checked to ensure that no materials designated as privileged or otherwise protected against disclosure were accidentally included in the production. Quality checking can include:
 1. Random spot checking of the production and re-review of particular records;
 2. Re-running keyword searches over the production similar to those used in creating the “potentially privileged” set. Documents returned by the search can be re-examined to ensure they are not privileged; and
 3. Running data format or pattern searches over the production to find potentially unredacted, but protected, personal information (e.g., a search for “###-##-####” to detect Social Security numbers).
- Where structured data is being produced, samples of the production set should be reviewed in the format in which it is being produced, to ensure that the correct data fields are present and that any fields that were redacted are not included in the production.

As a final note, to the extent possible, the legal team should organize its review in advance so that objective data (e.g., authors, recipients, dates, etc.) can be automatically populated into privilege logs. Contemporary review tools can support this process to a considerable extent, greatly simplifying the process of generating privilege logs. Such information, however, should be reviewed for accuracy and completeness.

CONCLUSION

“The future is already here – it’s just not evenly distributed.”¹⁰¹

Tremendous growth in the volume and complexity of ESI, and increasing scrutiny of the e-discovery process by opposing parties and the courts, bring with them increased risk of spoliation sanctions that compel the legal profession to implement best practices to achieve greater quality in all phases of the e-discovery process, including in its collection, culling, review, and production phases. In today’s legal environment, using project management, measures of quality, and statistical sampling, are some of the ways in which to adopt lean, efficient and smart business practices. Put another way: just as Moneyball¹⁰² demonstrated the value of applying new statistical measures to assess baseball talent (even while running counter to “tried and true practices” based on intuition and culture), this

101 William Gibson, quoted in THE ECONOMIST (Dec. 4, 2003).

102 See Michael Lewis, MONEYBALL: THE ART OF WINNING AN UNFAIR GAME (2003). For a good summary of the book’s thesis, see <http://en.wikipedia.org/wiki/Moneyball> (last accessed Nov. 1, 2013).

commentary posits that legal practice needs to incorporate the best thinking from other disciplines – even if lawyers need to learn new techniques, and even if some of the “mystique” of legal work is left behind.¹⁰³

In the end, cost-conscious firms, organizations, and institutions of all types that are intent on best practices and overburdened judges will demand that parties undertake new ways of thinking about how to solve e-discovery problems – including employing better project management and better measures of quality to achieve optimum results, as outlined in this commentary. The technical and project management-oriented quality processes discussed here (as well as others) should be incorporated into every litigator’s continuing education and daily practice.

These processes also dovetail with, and support, *The Sedona Conference Cooperation Proclamation* – which calls for incorporation of the best thinking of “disciplines outside the law” to achieve the goal of the “just, speedy, and inexpensive determination of every action and proceeding.” In the end, striving to attain a quality outcome in the conduct of litigation is consistent with the highest ethical calling of the legal profession.

103 See also Ian Ayres, *SUPERCRUNCHERS: WHY THINKING BY NUMBERS IS THE NEW WAY TO BE SMART* (Bantam Dell 2007). Ayres, a Yale Law School professor and econometrician, writes:

We are in a historic moment of horse-versus-locomotive competition, where intuitive and experiential expertise is losing out time and time again to number crunching. In the old days many decisions were simply based on some mixture of experience and intuition. Experts were ordained because of their decades of individual trial and error experience. We could trust that they knew the best way to do things, because they’d done it hundreds of times in the past. Experiential experts had survived and thrived. If you wanted to know what to do, you’d ask the grey-hairs. Now something is changing. Business and government professionals are relying more and more on databases to guide their decisions.

For a general overview of the growing field of analytics and business intelligence, see Thomas H. Davenport & Jeanne G. Harris, *COMPETING ON ANALYTICS: THE NEW SCIENCE OF WINNING* (2007).

APPENDIX: SAMPLING 101 FOR THE E-DISCOVERY LAWYER

The basic concept of sampling certain items from a population to better understand the characteristics of the whole population is simple and straightforward. Sampling is a familiar concept that is accepted by most people, including those tasked with reviewing electronic documents for responsiveness, and privilege. The application of sampling that most people are probably familiar with is with polling in national elections. Pollsters sample only a very, very small percentage of the voting public, but must take great care so as to ensure that the small number they are polling constitutes a genuinely representative sample of the entire population. On the basis of such a small but representative sample, the pollsters can predict:

- What percentage of voters would vote for a particular candidate;
- What the margin of error of the poll is (in statistical parlance, this is known as the “confidence interval”); and
- How certain they are that this result would hold up – within the margin of error – when applied to the population at large. (This is known as the “confidence level.”)

But mention “statistical sampling” in most other contexts and there is such apprehension about the process and the math involved, that, at least until very recently, statistical sampling was rarely used or cited in assessing the quality and consistency of the electronic document review process. Ultimately, lawyers are left in an unfortunate situation, since the document review process is well-suited to the application of statistical sampling to improve quality and reduce cost.

This section seeks to demystify the basics of statistical sampling and to explain how it can be effectively applied to a typical, large electronic document review process.

Some Basic Terms*Sampling: Judgmental (Nonprobabilistic) Versus Statistical*

Sampling involves the use of a subset of a population to estimate one or more characteristics of the whole population. Probability, statistical, or random sampling, is a sampling technique in which the probability of getting any particular sample may be calculated. “Judgmental” or “nonprobability sampling” does not meet this criterion – thus, judgmental sampling techniques cannot be used to infer from the sample to the general population. Any generalizations obtained from a nonprobability sample must be filtered through one’s knowledge of the topic being studied. Performing nonprobability sampling is generally considered less expensive than doing probability sampling, but the results are of more limited value.

Examples of Nonprobability Sampling Include:

Judgmental sampling: Sampling performed on a sample set that was selected based on the judgment of the person doing the sampling. For example, a researcher chooses the sample based on who or what they think would be appropriate for the study, as in conducting an isolated case study of just one group or making choices based on relationship proximity, not science. A common example in the e-discovery context would be keyword searching itself,

which is a technique universally used by lawyers and legal professionals to cull a set of data, based on the *a priori* judgment of those selecting the keyword terms.

Examples of Statistical Sampling:

“Random Sampling” A sample is a subset chosen from a population for investigation. A random sample is one chosen by a method involving an unpredictable component. Random sampling can also refer to taking a number of independent observations from the same probability distribution without involving any real population. A probability sample is one in which each item has a known (and equal) probability of being in the sample.

The sample will usually not be completely representative of the population from which it was drawn – this random variation in the results is known as “sampling error.” In the case of random samples, mathematical theory is available to assess the sampling error. Thus, estimates obtained from random samples can be accompanied by measures of the uncertainty associated with the estimate. This can take the form of a standard error, or if the sample is large enough for the central limit theorem to take effect, confidence intervals may be calculated.

A simple random sample is selected so that every possible item has an equal chance of being selected from the population.

A “self-weighting sample,” also known as an “Equal Probability of Selection Method (EPSEM) sample,” is one in which every individual, or object, in the population of interest has an equal opportunity of being selected for the sample. Simple random samples are self-weighting.

“Stratified Sampling” involves selecting independent samples from a number of subpopulations, groups, or strata within the population. Great gains in efficiency are sometimes possible from judicious stratification. An example of this would be varying the sampling percentage based on document type.

“Cluster Sampling” involves selecting the sample units in groups. For example, a sample of telephone calls may be collected by first taking a collection of telephone lines and collecting all the calls on the sampled lines. The analysis of cluster samples must take into account the intra-cluster correlation, which reflects the fact that units in the same cluster are likely to be more similar than two units picked at random.

Additional Common Sampling Terminology:

“Blind Sample” is a selected sample whose composition is unknown except to the person submitting it. This type of sample is used to test the validity of the measurement process.

“Acceptance Sampling” is a statistical procedure used for accepting or rejecting a batch of merchandise or documents. Acceptance sampling involves determining the maximum number of defects that may be discovered in a sample before the entire batch is rejected.

“Confidence Interval” is the range that contains the true population value estimate a specified percentage of the time, if repeated sampling of the population were to be performed. For example, a 95% confidence interval is a range that contains the true population value 95% of the time. A smaller range indicates an estimate that is more precise. Small sample sizes or cells with low numbers generate less precise estimates and will have wider confidence intervals.

“Confidence Level” is a statistical measure of the number of times out of 100 that the results measured will occur within a specified range. That is, a confidence level of 95% indicates that the result of an action will likely meet the expectations of the observer 95% of the time.

Statistical Sampling Basics

There are three key factors that determine the reliability and precision of the inferences that can be drawn about a population based on a statistical sample:

1. The **randomness** of the sample selection;
2. The **variability** of the sample results; and
3. The **sample size**.

One expects a properly-drawn, random sample to be representative of the population from which it comes, but any estimate based on the sample will naturally differ to some extent from a corresponding measurement based on the whole population. One expresses the likely extent of difference by presenting a confidence interval, i.e., a 95% confidence interval, around the estimated value. One will have confidence that the true value computed from the whole population would fall within the confidence interval 95% of the time. That is, if one draws samples over and over and computes estimates and intervals in the same way, one will be capturing the true value within the confidence interval 95% of the time.

The variability of the sample results affect how precise one can be in describing the true population values. This sampling variability is summarized by the quantity called the “standard error,” which is used in the construction of confidence intervals. One may say, “the population contains X% items, + or – Y%,” when X is the estimated value and Y is its standard error. In many cases, a confidence interval stretches about two standard errors above and below the estimate.

A simple sampling example might help to illustrate these concepts more effectively. If one were to reach into a backpack full of coins, mix them up thoroughly, and then pull out a single handful of all pennies, one might naturally believe that the bag contains all pennies based on the single handful that had been directly examined. However, a confidence interval for the proportion of pennies would stretch from zero percent to some low, but non-zero percent, depending on how many coins there were in the handful. The thorough mixing of the bag’s contents helps to achieve the randomness of the sample selection. However, the bag might well contain 2% dimes, and a handful-sized sample might happen by chance not to scoop up any of the dimes. The width of the confidence interval expresses this uncertainty.

This example is simplistic and it is easy to imagine dumping out the contents of the entire bag to confirm the inference based on the relatively random sample of one handful from one bag. But how sure can one really be of the remaining contents of the bag based on looking at a single handful of pennies? There would always be some doubt about the remaining contents of the bag. What if the handful of coins included several different denominations instead of all pennies? What could be said about the bag's contents then? And what if there were hundreds or thousands of bags full of coins? And, what if samples taken from a number of them yielded an assortment of coins instead of all one denomination? Making an inference about the contents of all of the bags, and perhaps the value of the all the coins in all of the bags combined, becomes a more challenging problem, but one that has been largely solved by statistics, as long as the objective is to get a reasonable estimated value for the population and not necessarily the exact value.

The mathematical formulae used in statistical sampling provide reliable, quantifiable estimates for making statements or inferences about a population based on the sample results. And, the less variability found in the sample results, the narrower a confidence interval can be, keeping the same level of confidence. Together, the variability of the sample results and the size of the sample are the two critical determinants of how wide a 95% confidence interval will be. With a larger sample, we can have a tighter interval within which we are confident of capturing the true value. But uncertainty decreases slowly with sample size. We must quadruple the size of the sample to halve the width of a confidence interval.

Since the basic concept of sampling is relatively simple, why is the application of sampling in the realm of document review so challenging to implement? Perhaps it is because of the perceived high level of risk of getting the "wrong" answer about any single item or document in the population. Building on the previous example, it seems that the approach of dumping out all the bags and examining all their contents would lower the risk of missing a few gold coins that might not be predicted or detected by sampling from the population. When performing document review for a legal proceeding, the attorneys involved will typically take an analogous approach by gathering all the documents and files from selected individuals and reviewing all that contain certain keywords. This approach should lower the absolute risk of missing a "smoking gun" or "privileged" document that could dramatically affect the outcome of the matter involved.

Since sampling can typically achieve only a "reasonable" rather than an "absolute" level of confidence about the entire population, it is often considered to be unsuitable for determining whether a group of documents contains any that are privileged or responsive. Presumably, the only way to determine whether documents are privileged or responsive is to look at them all and have the reviewers identify those that are privileged or responsive. But because the reviewers are only people, and people make mistakes, it makes sense to test their work for accuracy and consistency.

Sampling can be a very efficient method of determining whether or not reviewers have achieved the necessary and acceptable level of quality and consistency in their work. There are two distinct approaches to sampling to find the proportion of correctly marked documents in the population. The first involves process sampling for acceptance, based on an approach that examines or inspects documents selected from in-process batches. The second, quality-control testing, is performed by selecting a random sample from the entire population. Both of these methods can, in theory, be applied to a document review project.

The number of items to be tested or inspected is referred to as the sample size. The overall number of sample items selected from the entire review population for quality-control testing will typically be less than the sum of the individual samples drawn to test in-process review batches. However, the benefit of testing individual review batches is that the review process can be adjusted to improve quality as the review is performed, thereby increasing the likelihood that the overall population will meet the established “acceptable quality limits” when the process is completed.

THE SEDONA
CONFERENCE GLOSSARY:
E-DISCOVERY & DIGITAL INFORMATION
MANAGEMENT* (FOURTH EDITION)

*A Project of The Sedona Conference Working Group on
Electronic Document Retention & Production (WG1).*

Editors:

Sherry B. Harris, Crowley Law Office
Paul H. McVoy, Milberg LLP

RFP+ eDiscovery Vendor Panel

(see www.thesedonaconference.org for a listing of the
RFP+ eDiscovery Vendor Panel members)

Contributing Editor:

Matt K. Mulder, KPMG

Thanks go to all who participated in the dialogue that led to this Commentary. We thank all of our Working Group Series Sustaining and Annual Sponsors, whose support is essential to our ability to develop Working Group Series publications. For a listing of our sponsors just click on the “Sponsors” Navigation bar on the homepage of our website.

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.*

PREFACE

The Sedona Conference Glossary is published as a tool to assist in the understanding and discussion of electronic discovery and electronic information management issues. It is not intended to be an all-encompassing replacement of existing technical glossaries published by ARMA International (www.arma.org), American National Standards Institute (www.ansi.org), International Organization for Standardization (www.iso.org), U.S. National Archives & Records Administration (www.archives.gov) and other professional organizations.

The Sedona Conference acknowledges the contributions of all the Working Group 1 and RFP+ eDiscovery Vendor Panel members who reviewed and commented on drafts of this Fourth Edition, and thanks all of our Working Group Series Sponsors, whose support is essential to the ability of The Sedona Conference to develop Working Group Series publications such as this. For a complete listing of our Sponsors, click on the “Sponsors” navigation bar on the home page of our website at <http://www.thosedonaconference.org>.

As with all of our publications, your comments are welcome. Please forward them to me at cww@sedonaconference.org.

Craig Weinlein
Executive Director
The Sedona Conference
Phoenix, AZ
USA
April 2014

30(b)(6): A short hand reference to Rule 30(b)(6) of the Federal Rules of Civil Procedure, under which a corporation, partnership, association or governmental agency is subject to the deposition process, and required to provide one or more witnesses to “testify as to matters known or reasonably available to the organization” on the topics requested by the deposition notice. Sometimes the 30(b)(6) topics concern the discovery process itself, including procedures for preservation, collection, chain of custody, processing, review, and production.

Ablate: To burn laser-readable “pits” into the recorded layer of optical disks, DVD-ROMs and CD-ROMs.

Ablative: Unalterable data. See Ablate.

Access Control List (ACL): A security group comprised of individual users or user groups that is used to grant similar permissions to a program, database or other security controlled environment.

Active Data: Information residing on the direct access storage media (disk drives or servers) that is readily visible to the operating system and/or application software with which it was created. It is immediately accessible to users without restoration or reconstruction.

Active Records: Records related to current, ongoing or in-process activities referred to on a regular basis to respond to day-to-day operational requirements. See Inactive Record.

Address: A structured format for identifying the specific location or routing detail for information on a network or the Internet. These include email addresses Simple Mail Transfer Protocol (SMTP), Internet Protocol (IP) addresses and Uniform Resource Locators (URLs), commonly known as Web addresses.

Adware: See Spyware.

Agent: A program running on a computer that performs as instructed by a central control point to track file and operating system events and takes directed actions, such as transferring a file or deleting a local copy of a file, in response to such events.

AIIM: The Association for Information and Image Management, <http://www.aiim.org>, that focuses on Enterprise Content Management (ECM).

Algorithm: With regard to electronic discovery, computer script that is designed to analyze data patterns using mathematical formulas, and is commonly used to group or find similar documents based on common mathematical scores.

Alphanumeric: Characters composed of letters, numbers (and sometimes non-control characters, such as @, #, \$). Excludes control characters.

Ambient Data: See Data.

American National Standards Institute (ANSI): <http://www.ansi.org>—a private, non-profit organization that administers and coordinates the U.S. voluntary standardization and conformity assessment system.

American Standard Code for Information Interchange (ASCII, pronounced “ask-ee”):

A non-proprietary text format built on a set of 128 (or 255 for extended ASCII) alphanumeric and control characters. Documents in ASCII format consist of only text with no formatting and can be read by most computer systems.

Analog: Data in an analog format is represented by continuously variable, measurable, physical quantities such as voltage, amplitude or frequency. Analog is the opposite of digital.

Annotation: The changes, additions or editorial comments made or applicable to a document – usually an electronic image file – using electronic sticky notes, highlighter or other electronic tools. Annotations should be overlaid and not change the original document.

Aperture Card: An IBM punch card with a window that holds a 35mm frame of microfilm. Indexing information is punched in the card.

Applet: A program designed as an add-on to another program, allowing greater functionality for a specific purpose other than for what the original program was designed.

Appliance: A prepackaged piece of hardware and software designed to perform a specific function on a computer network, for example, a firewall.

Application: Software that is programmed for one or more specific uses or purposes. The term is commonly used in place of “program” or “software.” Applications may be designed for individual users, for example a word processing program or for multiple users, as in an accounting application used by many users at the same time.

Application Programming Interface (API): The specifications designed into a program that allows interaction with other programs. See Mail Application Programming Interface (MAPI).

Application Service Provider (ASP): An Internet-based organization hosting applications on its own servers within its own facilities. Customers license the application and access it through a browser over the Internet or via some other network. See Software as a Solution (SaaS).

Architecture: Refers to the hardware, software or combination of hardware and software comprising a computer system or network. “Open architecture” describes computer and network components that are more readily interconnected and interoperable. “Closed architecture” describes components that are less readily interconnected and interoperable.

Archival Data: Information an organization maintains for long-term storage and record keeping purposes but which is not immediately accessible to the user of a computer system. Archival data may be written to removable media such as a CD, magneto-optical media, tape or other electronic storage device or may be maintained on system hard drives. Some systems allow users to retrieve archival data directly while other systems require the intervention of an IT professional.

Archive, Electronic: Long-term repositories for the storage of records. Electronic archives preserve the content, prevent or track alterations, and control access to electronic records.

ARMA International: A not-for-profit association and recognized authority on managing records and information, both paper and electronic, <http://www.arma.org>.

Artificial Intelligence (AI): A subfield of computer science focused on the development of intelligence in machines so that the machines can react and adapt to their environment and the unknown. AI is the capability of a device to perform functions that are normally associated with human intelligence, such as reasoning and optimization through experience. It attempts to approximate the results of human reasoning by organizing and manipulating factual and heuristic knowledge. Areas of AI activity include expert systems, natural language understanding, speech recognition, vision, and robotics.

Aspect Ratio: The relationship of the height to the width of any image. The aspect ratio of an image must be maintained to prevent distortion.

Association for Computing Machinery (ACM): Professional association for computer professionals with a number of resources, including a special interest group on search and retrieval. See <http://www.acm.org>.

Attachment: A record or file associated with another record for the purpose of retention, transfer, processing, review, production and routine records management. There may be multiple attachments associated with a single “parent” or “master” record. In many records and information management programs or in a litigation context, the attachments and associated record(s) may be managed and processed as a single unit. In common use, this term often refers to a file (or files) associated with an email for retention and storage as a single Message Unit. See Document (or Document Family), Message Unit and Unitization.

Attribute: A specific property of a file such as location, size or type. The term attribute is sometimes used synonymously with “data element” or “property.”

Audio-Video Interleave (AVI): A Microsoft standard for Windows animation files that interleaves audio and video to provide medium quality multimedia.

Audit Log or Audit Trail: An automated or manual set of chronological records of system activities that may enable the reconstruction and examination of a sequence of events and/or changes in an event.

Author or Originator: The person, office or designated position responsible for an item’s creation or issuance. In the case of a document in the form of a letter, the author or originator is usually indicated on the letterhead or by signature. In some cases, a software application producing a document may capture the author’s identity and associate it with the document. For records management purposes, the author or originator may be designated as a person, official title, office symbol or code.

Auto-Delete: The use of technology to run predefined rules at scheduled intervals to delete or otherwise manage electronically stored information. May also be referred to as a janitor program or system cleanup.

Avatar: A graphical representation of a user in a shared virtual reality, such as Web forums or chat rooms.

Backbone: The top level of a hierarchical network. It is the main channel along which data is transferred.

Backup: The process of creating a copy of active data as a precaution against the loss or damage of the original data. The process is usually automated on a regular schedule, which can include the automatic expiration of older versions. The term is also used to refer to the Electronically Stored Information itself, as in, “a backup of the email server exists.” Backups can be made to any type of storage, including portable media, CDs, DVDs, data tapes or hard drives – also known as a full backup. See Differential Backup and Incremental Backup.

Backup Data: A copy of Electronically Stored Information that serves as a source for recovery in the event of a system problem or disaster. See Backup.

Backup Tape: Magnetic tape used to store copies of Electronically Stored Information, for use when restoration or recovery is required. The creation of backup tapes is made using any of a number of specific software programs and usually involves varying degrees of compression.

Backup Tape Rotation or Recycling: The process whereby an organization’s backups are overwritten with new data, usually on an automated schedule which should be determined by IT in consultation with records management and legal. For example, the use of nightly backup tapes for each day of the week – with the daily backup tape for a particular day being overwritten on the same day the following week.

Bandwidth: The amount of data a network connection can accommodate in a given period of time. Bandwidth is usually stated in kilobits per second (kbps), megabits per second (mps) or gigabytes per second (gps).

Bar Code: A small pattern of vertical lines or dots that can be read by a laser or an optical scanner. In records management and electronic discovery, bar codes may be affixed to specific records for indexing, tracking and retrieval purposes.

Basic Input Output System (BIOS): The set of user-independent computer instructions stored in a computer’s ROM, immediately available to the computer when the computer is turned on. BIOS information provides the code necessary to control the keyboard, display screen, disk drives and communication ports in addition to handling certain miscellaneous functions.

Batch File: A set of commands written for a specific program to complete a discrete series of actions, for example, renaming a series of files en masse.

Batch Processing: The processing of multiple sets of Electronically Stored Information at one time. See Processing Data.

Bates Number: Sequential numbering system used to identify individual pages of documents where each page or file is assigned a unique number. Often used in conjunction with a suffix or prefix to identify a producing party, the litigation or other relevant information. See Beginning Document Number and Production Number.

Bayesian Search: An advanced search that utilizes the statistical approach developed by Thomas Bayes, an 18th century mathematician and clergyman. Bayes published a theorem that describes how to calculate conditional probabilities from the combinations of observed events and prior probabilities. Many information retrieval systems implicitly or explicitly use Bayes’ probability rules to compute the likelihood that a document is relevant to a

query. For a more thorough discussion, see The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, available for download at <https://thesedonaconference.org/download-pub/3669>.

Beginning Document Number or BegDoc#: A unique number identifying the first page of a document or a number assigned to identify a native file.

Big Data: Enormous volumes of data, often distributed and loosely structured, that may be challenging to process with traditional technology solutions.

Bibliographic Coding: Manually recording objective information from documents such as date, authors, recipients, carbon copies, blind copies, and associating the information with a specific document. See Indexing and Coding.

Binary: The Base 2 numbering system used in digital computing that represents all numbers using combinations of zero and one.

Bit: Binary digit – the smallest unit of computer data. A bit consists of either 0 or 1. There are eight bits in a byte. See Byte.

Bitmap: A file format that contains information on the placement and color of individual bits used to convey images composed of individual bits (pixels), for which the system file extension is .bmp.

Bit Stream Backup: A sector-by-sector/bit-by-bit copy of a hard drive; an exact copy of a hard drive, preserving all latent data in addition to the files and directory structures. See Forensic Copy.

Cited: *Nucor Corp. v. Bell*, 2008 WL 4442571, D.S.C., January 11, 2008 (C/A No. 2:06-CV-02972-DCN.), quoting The Sedona Conference *Glossary: E-Discovery & Digital Information Management*, May 2005 First Edition (“A Bit Stream Back-up is an exact copy of a hard drive, preserving all latent data in addition to the files and directory structures. Bit Stream Back-up may be created using applications such as EnCase.”).

Cited: *U.S. v. Saboonchi*, 2014 WL 1364765 (D. Mar. Apr. 7, 2014).

Bitonal: A bitonal image uses only black and white.

Bits Per Inch (BPI): A unit of measure of data densities in disk and magnetic tape systems.

Bits Per Second (BPS): A measurement of the rate of data transfer. See Bandwidth.

Blowback: The term for printing Electronically Stored Information to hardcopy.

BMP: See Bitmap.

Bookmark: A link to another location, either within the current file or location or to an external location, like a specific address on the internet.

Boolean Search: Boolean searches use keywords and logical operators such as “and,” “or,” and “not” to include or exclude terms from a search, and thus produce broader or narrower search results. See Natural Language Search.

Boot Sector/Record: See Master Boot Sector/Record and Volume Boot Sector/Record.

Broadband: Commonly used in the context of high bandwidth Internet access made available through a variety of quickly evolving technologies.

Brontobyte: 1,024 yottabytes. See Byte.

Browser: An application used to view and navigate the World Wide Web and other Internet resources.

Bulletin Board System (BBS): A computer system or service that users access to participate in electronic discussion groups, post messages and/or download files.

Burn: The process of moving or copying data to portable media such as a CD or DVD.

Bus: A parallel circuit that connects the major components of a computer, allowing the transfer of electric impulses from one connected component to any other.

Byte (Binary Term): A basic measurement of most computer data consisting of 8 bits. Computer storage capacity is generally measured in bytes. Although characters are stored in bytes, a few bytes are of little use for storing a large amount of data. Therefore, storage is measured in larger increments of bytes. See Kilobyte, Megabyte, Gigabyte, Terabyte, Petabyte, Exabyte, Zettabyte, Yottabyte, Brontobyte and Geopbyte (listed here in order of increasing volume).

Cache: A dedicated, temporary, high speed storage location that can be used to store frequently-used data for quick user access, allowing applications to run more quickly.

Case De-Duplication: Eliminates duplicates to retain only one copy of each document per case. For example, if an identical document resides with three custodians, only the first custodian's copy will be saved. Also known as Cross Custodial De-Duplication, Global De-Duplication or Horizontal De-Duplication.

Catalog: See Index.

CCITT Group 4: A lossless compression technique/format that reduces the size of a file, generally about 5:1 over RLE and 40:1 over bitmap. CCITT Group 4 compression may only be used for bitonal images.

CD-R, CD+R (Compact Disk Recordable): See Compact Disk.

CD-RW (Compact Disk Re-Writable): See Compact Disk.

CD-ROM (Compact Disk Read-Only Memory): See Compact Disk.

Cellular Digital Packet Data (CDPD): A data communication standard utilizing the unused capacity of cellular voice providers to transfer data.

Central Processing Unit (CPU): The primary silicon chip that runs a computer's operating system and application software. It performs a computer's essential mathematical functions and controls essential operations.

Certificate: An electronic affidavit vouching for the identity of the transmitter. See Digital Certificate, Digital Signature, Public Key Infrastructure (PKI) Digital Signature.

Chain of Custody: Documentation regarding the possession, movement, handling and location of evidence from the time it is identified to the time it is presented in court or otherwise transferred or submitted; necessary both to establish admissibility and authenticity; important to help mitigate risk of spoliation claims.

Checksum: A value calculated on a set of data as a means of verifying its authenticity to a copy of the same set of data, usually used to ensure data was not corrupted during storage or transmission.

Child: See Document.

Clawback Agreement: An agreement outlining procedures to be followed if documents or Electronically Stored Information are inadvertently produced; typically used to protect against the waiver of privilege.

Client: (1) In a network, a computer that can obtain information and access applications on a server; (2) an application on a hard drive that relies on a server to perform some operations. See Thin Client.

Client Server: An architecture whereby a computer system consists of one or more server computers and numerous client computers (workstations). The system is functionally distributed across several nodes on a network and is typified by a high degree of parallel processing across distributed nodes. With client-server architecture, CPU intensive processes (such as searching and indexing) are completed on the server, while image viewing and OCR occur on the client. This dramatically reduces network data traffic and insulates the database from workstation interruptions.

Clipboard: A holding area in a computer's memory that temporarily stores information copied or cut from a document or file.

Cloud Computing: “[A] model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”
<http://csrc.nist.gov/publications/PubsSPs.html#800-145> (last visited January 29, 2013). For further discussion see the cited NIST publication SP800-145.pdf.

Cluster (File): The smallest unit of storage space that can be allocated to store a file on operating systems. Windows and DOS organize hard disks based on clusters (also known as allocation units), which consist of one or more contiguous sectors. Disks using smaller cluster sizes waste less space and store information more efficiently.

Cluster (System): A collection of individual computers that appear as a single logical unit. Also referred to as matrix or grid systems.

Cluster Bitmap: Used in NTFS (New Technology File System) to keep track of the status (free or used) of clusters on the hard drive. See NTFS.

Clustering: See Data Categorization.

Coding: Automated or human process by which specific information is captured from documents. Coding may be structured (limited to the selection of one of a finite number of choices) or unstructured (a narrative comment about a document). See Indexing, Verbatim Coding, Bibliographic Coding, Level Coding and Subjective Coding.

Color Graphics Adapter (CGA): See Video Graphics Adapter (VGA).

Comma Separated Value (CSV): A text file used for the transmission of fielded data that separates data fields with a comma and typically encloses data in quotation marks.

Commercial Off-the-Shelf (COTS): Hardware or software products that are commercially manufactured, ready-made and available for use by the general public without the need for customization.

Compact Disk (CD): A type of optical disk storage media; compact disks come in a variety of formats. These formats include CD-ROM (CD Read-Only Memory) – read-only; CD-R or CD+R (CD Recordable) – can be written to once and are then read-only; and CD-RW (CD Re-Writable) – can be written to multiple times.

Compliance Search: The identification of and search for relevant terms and/or parties in response to a discovery request.

Compound Document: A file that contains multiple files, often from different applications, by embedding objects or linked data; multiple elements may be included, such as images, text, animation or hypertext. See Container File and OLE.

Compression: The reduction in the size of a source file or files with the use of variety of algorithms, depending on the software being used. Algorithms approach the task in a variety of ways, generally eliminates redundant information or by predicting where changes are likely to occur.

Compression Ratio: The ratio of the size of an uncompressed file to a compressed file, e.g., with a 10:1 compression ratio. Example: a 10 KB file can be compressed to 1 KB.

Computer: Includes but is not limited to network servers, desktops, laptops, notebook computers, mainframes and PDAs (personal digital assistants).

Computer Aided Design (CAD): The use of a wide range of computer-based tools that assist engineers, architects and other design professionals in their design activities.

Computer Aided or Assisted Review: See Technology-Assisted Review.

Computer Client or Client: A computer or program that requests a service of another computer system. A workstation requesting the contents of a file from a file server is a client of the file server. See Thin Client. Also commonly used as synonymous with an email application, by reference to the Email Client. See Thin Client.

Computer Forensics: The use of specialized techniques for recovery, authentication and analysis of electronic data when an investigation or litigation involves issues relating to reconstruction of computer usage, examination of residual data, authentication of data by technical analysis or explanation of technical features of data and computer usage. Computer forensics requires specialized expertise that goes beyond normal data collection

and preservation techniques available to end-users or system support personnel and generally requires strict adherence to chain-of-custody protocols. See Forensics and Forensic Copy.

Computer Output to Laser Disk (COLD): A computer programming process that outputs electronic records and printed reports to laser disk instead of a printer.

Computer Output to Microfilm (COM): A process that outputs electronic records and computer generated reports to microfilm.

Concatenate: Generally, to add by linking or joining to form a chain or series; the process of linking two or more databases of similar structure to enable the user to search, use, or reference them as if they were a single database.

Concept Search: The method of search that uses word meanings and ideas, without the presence of a particular word or phrase, to locate Electronically Stored Information related to a desired concept. Word meanings can be derived from any of a number of sources, including dictionaries, thesauri, taxonomies, ontologies or computed mathematically from the context in which the words occur.

Conceptual Analytics: Using one or more of a number of mathematical algorithms or linguistic methodologies to analyze unstructured data by themes and ideas contained within the documents enabling the grouping or searching of documents or other unstructured data by their common themes or ideas.

Container File: A compressed file containing multiple files; used to minimize the size of the original files for storage and/or transporting. Examples include .zip, .pst and .nsf files. The file must be ripped or decompressed to determine volume, size, record count, etc. and to be processed for litigation review and production. Also see Decompression and Rip.

Cited: *Country Vintner of North Carolina, LLC v. E. & J. Gallo Winery, Inc.*, 718 F.3d 249 (4th Cir. Apr. 29, 2013).

Content Comparison: A method of de-duplication that compares file content or output (to image or paper) and ignores metadata. See De-Duplication.

Contextual Search: Using one of a number of mathematical algorithms or linguistic methodologies to enlarge search results to include not only exact term matches but also matches where terms are considered in context of how and where they frequently occur in a specific document collection or more general taxonomy. For example, a search for the term “diamond” may bring back documents related to baseball but with no reference to the word diamond because they frequently occur within the same documents and therefore have a logical association.

Cookie: A text file containing tracking information such as dates and times of website visits, deposited by a website onto a user’s computer or mobile device. The text file is accessed each time the website is visited by a specific user and updated with browsing and other information. The main purpose of cookies is to identify users and possibly prepare customized websites for them, including the personalization of advertising appearing on the websites.

Coordinated Universal Time (UTC): A high precision atomic time standard with uniform seconds defined by International time and leap seconds announced at regular intervals to compensate for the earth's slowing rotation and other discrepancies. Leap seconds allow UTC to closely track Universal time, a time standard based not on the uniform passage of seconds but on the earth's angular rotation. Time zones around the world are expressed as positive or negative offsets from UTC. Local time is UTC plus the time zone offset for that location plus an offset (typically +1) for daylight savings, if in effect. For example, 3:00 a.m. Mountain Standard Time = 10:00 UTC – 7. As the zero point reference, UTC is also referred to as Zulu time (Z). See Normalization.

Corrupted File: A file that has become damaged in some way, such as by a virus or by software or hardware failure, so that it is partially or completely unreadable by a computer.

Characters Per Inch (CPI): A description of the number of characters that are contained in an inch of backup tape.

Cross-Custodian De-Duplication: The suppression or removal of exact copies of files across multiple custodians for the purposes of minimizing the amount of data for review and/or production. See Case De-Duplication and De-Duplication.

Cryptography: Technique to scramble data to preserve confidentiality or authenticity.

Cull (verb): To remove or suppress from viewing, a document from a collection to be reviewed or produced. See Data Filtering and Harvesting.

Custodian: See Record Custodian and Record Owner.

Custodian De-Duplication: The removal or suppression of exact copies of a file found within a single custodian's data for the purposes of minimizing the amount of data for review and/or production. Also known as Vertical De-duplication. See De-Duplication.

Customer Relationship Management (CRM) Application: Computer program that helps manage communications with client and contains contact information.

Cyclical Redundancy Checking (CRC): Used in data communications to create a checksum character at the end of a data block to ensure integrity of data transmission and receipt. See Checksum.

Cylinder: The set of tracks on both sides of each platter in a hard drive that is located at the same head position. See Platter.

Data: Any information stored on a computer, whether created automatically by the computer, such as log files or created by a user, such as the information entered on a spreadsheet. See Active Data and Latent Data.

Data Categorization: The categorization and sorting of Electronically Stored Information – such as foldering by “concept,” content, subject, taxonomy, etc. – through the use of technology – such as search and retrieval software or artificial intelligence – to facilitate review and analysis.

Data Cell: An individual field of an individual record. For example, in a table containing information about all of a company's employees, information about employee Joe Smith is

stored in a single record, and information about his social security number is stored in an individual Data Cell. See Field.

Data Collection: See Harvesting.

Data Controller (as used in the EU Data Protection Act): The natural or legal person who alone or jointly with others determines the purposes for which and the manner in which any Personal Data are to be processed.

Data Element: A combination of characters or bytes referring to one separate piece of information, such as name, address or age.

Data Encryption Standard (DES): A form of private key encryption developed by IBM in the late 1970's.

Data Extraction: The process of parsing data, including the text of the file, from any electronic documents into separate metadata fields such as Date Created and Date Last Accessed.

Data Field: See Field.

Data Filtering: The process of identifying data based on specified parameters, such as date range, author(s), and/or keyword search terms, often used to segregate data for inclusion or exclusion in the document culling or review workflow.

Data Formats: The organization of information for display, storage or printing. Data is sometimes maintained in certain common formats so that it can be used by various programs, which may only work with data in a particular format, e.g. PDF, HTML. Also used by parties to refer to production specifications during the exchange of data during discovery.

Data Harvesting: See Harvesting.

Data Map: A document or visual representation that records the physical or network location and format of an organization's data. Information about the data can include where the data is stored, physically and virtually, in what format it is stored, backup procedures in place, how the Electronically Stored Information moves and is used throughout the organization, information about accessibility of the Electronically Stored Information, retention and lifecycle management practices and policies, and identity of records custodians.

Data Mining: The process of knowledge discovery in databases (structured data); often techniques for extracting information, summaries or reports from databases and data sets. In the context of electronic discovery, this term often refers to the processes used to analyze a collection of Electronically Stored Information to extract evidence for production or presentation in an investigation or in litigation. See Text Mining.

Data Processor (as used in the EU Data Protection Act): A natural or legal person (other than an employee of the Data Controller) who processes Personal Data on behalf of the Data Controller.

Data Set: A named or defined collection of data. See Production Data Set and Privilege Data Set.

Data Subject (as used in the EU Data Protection Act): An individual who is the subject of Personal Data.

Data Verification: Assessment of data to ensure it has not been modified from a prior version. The most common method of verification is hash coding by using industry accepted algorithms such as MD5, SHA1 or SHA2. See Digital Fingerprint, File Level Binary Comparison, and Hash Coding.

Database: A set of data elements consisting of at least one file or of a group of integrated files, usually stored in one location and made available to several users. The collection of information is organized into a predefined formatted structure and usually organized into fields of data that comprise individual records which are further grouped into data tables. Databases are sometimes classified according to their organizational approach, with the most prevalent approach being the relational database – a tabular database in which data is defined so that it can be reorganized and accessed in a number of different ways. Another popular organizational structure is the distributed database, which can be dispersed or replicated among different points in a network. Computer databases typically contain aggregations of data records or files, such as sales transactions, product catalogs and inventories, and customer profiles. For further discussion, see *The Sedona Conference Database Principles*, available for download at <http://www.thesedonaconference.org/download-pub/426>.

Database Management System (DBMS): A software system used to access and retrieve data stored in a database.

Date/Time Normalization: See Normalization.

Date Created: A common metadata field that contains the date a file was created on or moved to and the media where it currently resides.

Date Last Accessed: A common metadata field that contains the date a file was last accessed, meaning last opened or moved or even copied depending on the technology used to copy.

Date Last Modified: A common metadata field that contains the date a file was last changed either by a modification to the content or format, printed or changed by the automatic running of any macros that are executed upon the file being opened. The date last modified field does not normally reflect a change to a file's storage location or when the file was opened and read, and is thus often used as an electronic file date control field for discovery purposes.

Date Sent: A common metadata field that contains the date on which an email was sent.

Date Received: A common metadata field that contains the date on which an email was received.

Daubert or Daubert Challenge: *Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579, at 593-94 (1993), addresses the admission of scientific expert testimony to ensure that the testimony is reliable before considered for admission pursuant to Rule 702. The court assesses the testimony by analyzing the methodology and applicability of the expert's approach. Faced with a proffer of expert scientific testimony, the trial judge must determine first, pursuant to Rule 104(a), whether the expert is proposing to testify to (1) scientific

knowledge that (2) will assist the trier of fact to understand or determine a fact at issue. This involves preliminary assessment of whether the reasoning or methodology is scientifically valid and whether it can be applied to the facts at issue. Daubert suggests an open approach and provides a list of four potential factors: (1) whether the theory can be or has been tested; (2) whether the theory has been subjected to peer review or publication; (3) known or potential rate of error of that particular technique and the existence and maintenance of standards controlling the technique's operation; and (4) consideration of general acceptance within the scientific community.

Decompression: To expand or restore compressed data back to its original size and format. See Compression.

Decryption: Transformation of encrypted (or scrambled) data back to original form.

De-Duplication (“de-dupe”): The process of comparing electronic files or records based on their characteristics and removing, suppressing or marking exact duplicate files or records within the data set for the purposes of minimizing the amount of data for review and production. De-duplication is typically achieved by calculating a file or record's hash value using a mathematical algorithm. De-duplication can be selective, depending on the agreed-upon criteria. See Case De-Duplication, Content Comparison, Cross-Custodian De-Duplication, Custodian De-Duplication, Data Verification, Digital Fingerprint, File Level Binary Comparison, Hash Coding, Horizontal De-Duplication, Metadata Comparison, Near Duplicates.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

Defragment (“defrag”): Use of a computer utility to reorganize files so they are more physically contiguous on a hard drive or other storage medium, when the files or parts thereof have become fragmented and scattered in various locations within the storage medium in the course of normal computer operations. Used to optimize the operation of the computer, it will overwrite information in unallocated space. See Fragmentation.

Deleted Data: Information that is no longer readily accessible to a computer user due to the intentional or automatic deletion of the data. Deleted data may remain on storage media in whole or in part until overwritten or wiped. Even after the data itself has been wiped, directory entries, pointers or other information relating to the deleted data may remain on the computer. Soft deletions are data marked as deleted (and not generally available to the end-user after such marking) but not yet physically removed or overwritten. Soft-deleted data can be restored with complete integrity.

Deletion: The process whereby data is removed from active files and other data storage structures on computers and rendered more inaccessible except through the use of special data recovery tools designed to recover deleted data. Deletion occurs on several levels in modern computer systems: (a) File level deletion renders the file inaccessible to the operating system and normal application programs and marks the storage space occupied by the file's directory entry and contents as free and available to reuse for data storage; (b) Record level deletion occurs when a record is rendered inaccessible to a database management system (DBMS) (usually marking the record storage space as available for reuse by the DBMS, although in some cases the space is never reused until the database is compacted) and is also characteristic of many email systems; and (c) Byte level deletion occurs when text or other information is deleted from the file content (such as the deletion

of text from a word processing file); such deletion may render the deleted data inaccessible to the application intended to be used in processing the file, but may not actually remove the data from the file's content until a process such as compaction or rewriting of the file causes the deleted data to be overwritten.

De-NIST: The use of an automated filter program that screens files against the NIST list in order to remove files that are generally accepted to be system generated and have no substantive value in most instances. See NIST List.

De-shading: Removing shaded areas to render images more easily recognizable by OCR. De-shading software typically searches for areas with a regular pattern of tiny dots.

De-skewing: The process of straightening skewed (tilted) images. De-skewing is one of the image enhancements that can improve OCR accuracy. Documents often become skewed when scanned or faxed.

Desktop: Generally refers to the working area of the display on an individual PC.

Desktop Publishing (DTP): PC applications used to prepare direct print output or output suitable for printing presses.

De-speckling: Removing isolated speckles from an image file. Speckles often develop when a document is scanned or faxed. See Speckle.

Differential Backup: A method of backing up data that backs up data that is new or has been changed from that last full backup.

Digital: Information stored as a string of ones and zeros (numeric). Opposite of analog.

Digital Audio Disk (DAD): Another term for compact disk.

Digital Audio Tape (DAT): A magnetic tape generally used to record audio but can hold up to 40 gigabytes (or 60 CDs) of data if used for data storage. Has the disadvantage of being a serial access device. Often used for backup.

Digital Certificate: Electronic records that contain unique secure values used to decrypt information, especially information sent over a public network like the Internet. See Certificate, Digital Signature and Public Key Infrastructure (PKI) Digital Signature.

Digital Evidence Bag (DEB): A container file format used for electronic evidence to preserve and transfer evidence in an encrypted or protected form that prevents deliberate or accidental alteration. The secure wrapper provides metadata concerning the collection process and context for the contained data.

Digital Fingerprint: A fixed-length hash code that uniquely represents the binary content of a file. See Data Verification, File Level Binary Comparison, Hash Coding.

Digital Linear Tape (DLT): A type of magnetic computer tape used to copy data from an active system for purposes of archiving or disaster recovery.

Digital Millennium Copyright Act (DMCA): United States copyright law enacted to protect against copyright infringement of data, address rights and obligations of owners of

copyrighted material, and the rights and obligations of Internet service providers on whose systems the infringing material may reside.

Digital Rights Management (DRM): A program that controls access to, movement or duplication of protected data.

Digital Signature: A way to ensure the identity of the sender, utilizing public key cryptography and working in conjunction with certificates. See Certificate, Digital Certificate and Public Key Infrastructure (PKI) Digital Signature.

Digital to Analog Converter (DAC): Converts digital data to analog data.

Digital Video Disk or Digital Versatile Disk (DVD): A plastic disk, like a CD, on which data can be optically written and read. DVDs can hold more information and can support more data formats than CDs. Formats include: DVD-R or DVD+R (DVD Recordable) – written to once and are then read-only; and DVD-RW (DVD Re-Writable) – can be written to multiple times.

Digitize: The process of converting an analog value into a digital (numeric) representation. See Analog.

Directory: The organizational structure of a computer's file storage, usually arranged in a hierarchical series of folders and subfolders. Often simulated as a file folder tree.

Disaster Recovery Tapes: Portable magnetic storage media used to store data for backup purposes. See Backup Data, Backup Tape.

Discovery: The process of identifying, locating, preserving, securing, collecting, preparing, reviewing and producing facts, information and materials for the purpose of producing/obtaining evidence for utilization in the legal process. There are several ways to conduct discovery, the most common of which are interrogatories, requests for production of documents and depositions. See Electronic Discovery.

Disk: Round, flat storage media with layers of material that enable the recording of data.

Disk Mirroring: The ongoing process of making an exact copy of information from one location to another in real time and often used to protect data from a catastrophic hard disk failure or for long term data storage. See Mirror Image and Mirroring.

Disk Partition: A discrete section of a computer's hard drive that has been virtually separated from one or more other partitions on the same drive.

Diskwipe: Utility that overwrites existing data. Various utilities exist with varying degrees of efficiency – some wipe only named files or unallocated space of residual data, thus unsophisticated users who try to wipe evidence may leave behind files of which they are unaware.

Disposition: The final business action carried out on a record. This action generally is to destroy or archive the record. Electronic record disposition can include “soft deletions” (see Deletion), “hard deletions,” “hard deletions with overwrites,” “archive to long-term store,” “forward to organization,” and “copy to another media or format and delete (hard or soft).”

Distributed Data: Information belonging to an organization that resides on portable media and non-local devices such as remote offices, home computers, laptop computers, personal digital assistants (PDAs), wireless communication devices (e.g., Blackberry) and Internet repositories (including email hosted by Internet service providers or portals and websites). Distributed data also includes data held by third parties such as application service providers and business partners. Note: Information Technology organizations may define distributed data differently (for example, in some organizations distributed data includes any non-server-based data, including workstation disk drives).

Document (or Document Family): A collection of pages or files produced manually or by a software application, constituting a logical single communication of information, but consisting of more than a single stand-alone record. Examples include a fax cover, the faxed letter, and an attachment to the letter – the fax cover being the “Parent,” and the letter and attachment being a “Child.” See Attachment, Load File, Message Unit, and Unitization – Physical and Logical.

Cited: *Abu Dhabi Commercial Bank v. Morgan Stanley & Co. Inc.*, 2011 WL 3738979 (S.D.N.Y., August 18, 2011).

Document Date: Generally, the term used to describe the date the document was last modified or put in final form; applies equally to paper and electronic files. See Date Last Modified, Date Created, Date Last Accessed, Date Sent, Date Received.

Document Imaging Programs: Software used to scan paper documents and to store, manage, retrieve and distribute documents quickly and easily.

Document Type or Doc Type: A bibliographic coding field that captures the general classification of a document, i.e., whether the document is correspondence, memo, report, article and others.

DoD 5015: Department of Defense standard addressing records management.

Domain: A group of servers and computers connected via a network and administered centrally with common rules and permissions.

DOS: See Microsoft-Disk Operating System (MS-DOS).

Double Byte Language: See Unicode.

Download: To move data from a remote location to a local computer or network, usually over a network or the Internet; also used to indicate that data is being transmitted from a location to a location. See Upload.

Draft Record: A preliminary version of a record before it has been completed, finalized, accepted, validated or filed. Such records include working files and notes. Records and information management policies may provide for the destruction of draft records upon finalization, acceptance, validation or filing of the final or official version of the record. However, draft records generally must be retained if: (1) they are deemed to be subject to a legal hold; or (2) a specific law or regulation mandates their retention and policies should recognize such exceptions.

Drag-and-Drop: The movement of files by dragging them with the mouse and dropping them in another place.

Drive Geometry: A computer hard drive is made up of a number of rapidly rotating platters that have a set of read/write heads on both sides of each platter. Each platter is divided into a series of concentric rings called tracks. Each track is further divided into sections called sectors, and each sector is subdivided into bytes. Drive geometry refers to the number and positions of each of these structures.

Driver: A computer program that controls various hardware devices such as the keyboard, mouse or monitor and makes them operable with the computer.

Drop-Down Menu: A menu window that opens on-screen to display context-related options. Also called pop-up menu or pull-down menu.

Dynamic Data Exchange (DDE): A form of interprocess communications used by Microsoft Windows to support the exchange of commands and data between two simultaneously running applications.

Dynamic Random Access Memory (DRAM): A memory technology that is periodically refreshed or updated – as opposed to static RAM chips that do not require refreshing. The term is often used to refer to the memory chips themselves.

Early Data Assessment: The process of separating possibly relevant Electronically Stored Information from non-relevant Electronically Stored Information using both computer techniques, such as date filtering or advanced analytics, and human assisted logical determinations at the beginning of a case. This process may be used to reduce the volume of data collected for processing and review. Also known as Early Case Assessment (ECA).

Electronic Data Interchange (EDI): Eliminating forms altogether by encoding the data as close as possible to the point of the transaction; automated business information exchange.

Electronic Discovery (E-Discovery): The process of identifying, locating, preserving, collecting, preparing, reviewing, and producing Electronically Stored Information (ESI) in the context of the legal process. See Discovery.

Cited: *Gordon v. Kaleida Health*, 2013 WL 2250579 (W.D.N.Y. May 21, 2013).

Cited: *Hinterberger v. Catholic Health System, Inc.*, 2013 WL 2250603 (W.D.N.Y. May 21, 2013).

Electronic Document Management: The process of using a computer program to manage individual unstructured files, either those created electronically or scanned to digital form from paper. See Information Lifecycle Management.

Electronic Document Management System (EDMS): A system to electronically manage documents during all life cycles. See Electronic Document Management.

Electronic File Processing: See Processing Data.

Electronic Image: An individual page or pages of an electronic document that has been converted into a static format, for example PDF or TIFF. See PDF and TIFF.

Electronic Record: Information recorded in a form that requires a computer or other machine to process it.

Electronically Stored Information (ESI): As referenced in the United States Federal Rules of Civil Procedure, information that is stored electronically, regardless of the media or whether it is in the original format in which it was created, as opposed to stored in hard copy (i.e., on paper).

Email (Electronic Mail): An electronic means for sending, receiving and managing communications via a multitude of different structured data applications (email client software), such as Outlook or Lotus Notes or those often known as “webmail,” such as Gmail or Yahoo Mail. See Email Message.

Email Address: Unique value given to individual user accounts on a domain used to route email messages to the correct email recipient most often formatted as follows: user-ID@domain-name. See Email Message.

Email Archiving: A systematic approach to retaining and indexing email messages to provide centralized search and retrieval capabilities. See Journaling.

Email Client: See Email (Electronic Mail).

Email Message: A file created or received via an electronic mail system. Any attachments that may be transmitted with the email message are not part of the email message but are part of the Message Unit and Document Family.

Email String: An electronic conversation between two or more parties via email. Also referred to as an email thread. See Thread.

Email Store: File or database containing individual email messages. See Container File, Message Unit, OST, PST, and NSF.

Embedded Object: A file or piece of a file that is copied into another file, often retaining the utility of the original file’s application; for example, a part of a spreadsheet embedded into a word processing document that still allows for editing and calculations after being embedded. See Compound Document.

EML: File extension of a generic email message file.

Encapsulated PostScript (EPS): Uncompressed files for images, text and objects. Can only be printed on printers with PostScript drivers.

Encoding: To change or translate into code; to convert information into digital format. For software, encoding is used for video and audio references, like encoding analog format into digital or raw digital data into compressed format.

Encryption: A procedure that renders the contents of a message or file unreadable to anyone not authorized to read it; used to protect Electronically Stored Information being stored or transferred from one location to another.

Encryption Key: A data value that is used to encrypt and decrypt data. The number of bits in the encryption key is a rough measure of the encryption strength; generally, the more bits in the encryption key, the more difficult it is to break.

End Document Number or End Doc#: A common metadata field that contains the bates number of the last page of a document.

End of File (EOF): A distinctive code that uniquely marks the end of a data file.

Enhanced Parallel Port (EPP): See Port.

Enhanced Small Device Interface (ESDI): A defined, common electronic interface for transferring data between computers and peripherals, particularly disk drives.

Enhanced Titles: A bibliographic coding field that captures a meaningful/descriptive title for a document based on a reading of the document as opposed to a verbatim title lifted as it appears on the face of the document. See Verbatim Coding.

Enterprise Architecture: Framework of information systems and processes integrated across an organization. See Information Technology Infrastructure.

Enterprise Content Management (ECM): Management of an organization's unstructured Electronically Stored Information, regardless of where it exists, throughout the entire lifecycle of the Electronically Stored Information.

Ephemeral Data: Data that exists for a very brief, temporary period and is transitory in nature, such as data stored in RAM.

Erasable Optical Disk: A type of optical disk that can be erased and new Electronically Stored Information added; most optical disks are read only.

Ethernet: A common way of networking PCs to create a Local Area Network (LAN).

Evidentiary Image or Copy: See Forensic Copy.

Exabyte: – 1,024 petabytes (approximately one billion gigabytes). See Byte.

Exchange Server: A server running Microsoft Exchange messaging and collaboration software. It is widely used by enterprises using Microsoft infrastructure solutions. Among other things, Microsoft Exchange manages email, shared calendars, and tasks.

Expanded Data: See Decompression.

Export: The process of saving data or a subset of data in a format that can be used or imported by another system.

Extended Graphics Adapter (EGA): See Video Graphics Adapter (VGA).

Extended Partitions: If a computer hard drive has been divided into more than four partitions, extended partitions are created. Under such circumstances each extended partition contains a partition table in the first sector that describes how it is further subdivided. See Disk Partition.

Extensible Markup Language (XML): A software coding language specification developed by the W3C (World Wide Web Consortium – the Web development standards board). XML is a pared-down version of SGML, designed especially for Web documents. It allows designers to create their own customized tag, enabling the definition, transmission, validation, and interpretation of data between applications and between organizations.

Extranet: The portion of an intranet site that is accessible by users outside of a company or organization hosting the intranet. This type of access is often utilized in cases of joint defense, joint venture, and vendor client relationships.

False Negative: A result from a search that is not correct because it fails to indicate a match or hit where one exists.

False Positive: A result from a search that is not correct because it indicates a match or hit where there is none.

Fast Mode Parallel Port: See Port.

Federal Information Processing Standards (FIPS): Issued by the National Institute of Standards and Technology after approval by the Secretary of Commerce pursuant to Section 111(d) of the Federal Property and Administrative Services Act of 1949, as amended by the Computer Security Act of 1987, Public Law 100-235.

Fiber Optics: Transmitting information by sending light pulses over cables made from thin strands of glass.

Field (or Data Field): A defined area of a file or data table used to record an individual piece of standardized data, such as the author of a document, a recipient or the date of a document.

Field Mapping: The process of normalizing data to the structure of an existing database for purposes of loading the data so that data loads to the correct field after validating the data type is the same, for example mapping the data from a field called Date to an existing field in a database named DocDate.

Field Separator or Field Delimiter: A character in a text delimited file that separates the fields in an individual record. For example, the CSV format uses a comma as the field separator. See Text Delimited File.

File: A collection of related data or information stored as a unit under a specified name on storage medium.

File Allocation Table (FAT): An internal data table on hard drives that keeps track of where the files are stored. If a FAT is corrupt, a drive may be unusable, yet the data may be retrievable with forensics. See Cluster (File).

File Compression: See Compression.

File Extension: Many systems, including DOS and UNIX, allow a filename extension that consists of one or more characters following the proper filename. For example, image files are usually stored as .bmp, .gif, .jpg or .tiff. Audio files are often stored as .aud or .wav. There are a multitude of file extensions identifying file formats. The filename extension should indicate what type of file it is; however, users may change filename extensions to evade firewall restrictions or for other reasons. Therefore, file types should be identified at a binary level rather than relying on file extensions. To research file types, see <http://www.filext.com>. Different applications can often recognize only a predetermined selection of file types. See Format (noun).

File Format: The organization or characteristics of a file that determine with which software programs it can be used. See Format (noun).

File Header: See Header.

File Level Binary Comparison: Method of de-duplication using the digital fingerprint (hash) of a file to compare the individual content and location of bytes in one file against those of another file. See Data Verification, De-Duplication, Digital Fingerprint, Hash Coding.

File Plan: A document containing the identifying number, title, description, and disposition authority of files held or used in an office.

File Server: A computer that serves as a storage location for files on a network. File servers may be employed to store Electronically Stored Information, such as email, financial data or word processing information or to backup the network. See Server.

File Sharing: Providing access to files or programs to multiple users on a network.

File Signature: See Digital Signature.

File Slack: See Slack/Slack Space.

File System: The means by which an operating system or program organizes and keeps track of Electronically Stored Information in terms of logical structures and software routines to control access to stored Electronically Stored Information, including the structure in which the files are named, stored, and organized. The file system also tracks data when a user copies, moves or deletes a file or subdirectory.

File Table: See Master File Table (MFT).

File Transfer: The process of moving or transmitting a copy of a file from one location to another, as between two programs or from one computer to another.

File Transfer Protocol (FTP): An internet protocol that governs the transfer of files between computers over a network or the Internet. The terms FTP server or FTP site are commonly used to refer to a location to upload/download and exchange data, particularly in large volume.

Filename: The name used to identify a specific file in order to differentiate it from other files, typically comprised of a series of characters a dot and a file extension (e.g., sample.doc). See File Extension, Full Path.

Firewall: A set of related security programs and/or hardware that protect the resources of a private network from unauthorized access by users outside of an organization or user group. A firewall filters information to determine whether to forward the information toward its destination.

Filter (verb): See Data Filtering.

Flash Drive: A small removable data storage device that uses flash memory and connects via a USB port. Can be imaged and may contain residual data. Metadata detail may not be the equivalent of Electronically Stored Information maintained in more robust storage media.

Flash Memory: A type of computer memory used for storage of data to a physical disk by electrical impulses.

Flat File: A non-relational, text-based file (i.e., a word processing document).

Floppy Disk: A thin magnetic film disk housed in a protective sleeve used to copy and transport relatively small amounts of data.

Folder: See Directory.

Forensic Copy: An exact copy of an entire physical storage media (hard drive, CD-ROM, DVD-ROM, tape, etc.), including all active and residual data and unallocated or slack space on the media. Forensic copies are often called images or imaged copies. See Bit Stream Backup, Mirror Image.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

Forensics: The scientific examination and analysis of data held on, or retrieved from, a computer in such a way that the information can be used as evidence in a court of law. It may include the secure collection of computer data; the examination of suspect data to determine details such as origin and content; the presentation of computer based information to courts of law; and the application of a country's laws to computer practice. Forensics may involve recreating deleted or missing files from hard drives, validating dates and logged in authors/editors of documents, and certifying key elements of documents and/or hardware for legal purposes.

Form of Production: The specifications for the exchange of documents and/or data between parties during a legal dispute. Used to refer both to file format (e.g., native vs. imaged format with agreed-upon metadata and extracted text in a load file) and the media on which the documents are produced (paper vs. electronic). See Load File, Native Format.

Format (noun): The internal structure of a file, which defines the way it is stored and used. Specific applications may define unique formats for their data (e.g., "MS Word document file format"). Many files may only be viewed or printed using their originating application or an application designed to work with compatible formats. There are several common email formats, such as Outlook and Lotus Notes. Computer storage systems commonly identify files by a naming convention that denotes the format (and therefore the probable originating application). For example, DOC for Microsoft Word document files; XLS for Microsoft Excel spreadsheet files; TXT for text files; HTM for HyperText Markup Language (HTML) files such as web pages; PPT for Microsoft PowerPoint files; TIF for tiff images; PDF for Adobe images; etc. Users may choose alternate naming conventions, but this will likely affect how the files are treated by applications.

Format (verb): To make a drive ready to store data within a particular operating system. Erroneously thought to "wipe" drive. Typically, only overwrites the File Allocation Table, but not the actual files on the drive.

Forms Processing: A specialized imaging application designed for handling pre-printed forms. Forms processing systems often use high-end (or multiple) OCR engines and elaborate data validation routines to extract handwritten or poor quality print from forms that go into a database.

Fragmentation: The process by which parts of files are separately stored in different areas on a hard drive or removable disk in order to utilize available space. See Defragment.

Full Duplex: Data communications devices that allow full speed transmission between computers in both directions at the same time.

Full Path: A file location description that includes the drive, starting or root directory, all attached subdirectories and ending with the file or object name. Often referred to as the Path Name.

Full-Text Indexing: The extraction and compilation of text from a collection of ESI. Text is gathered both from the body of the data and selected metadata fields. See Index.

Full-Text Search: The ability to search an index of all the words in a collection of Electronically Stored Information for specific characters, words, numbers and/or combinations or patterns thereof in varying degrees of complexity.

Fuzzy Search: The method of searching an index that allows for one or more characters in the original search terms to be replaced by wild card characters so that a more broad range of data hits will be returned. For example, a fuzzy search for “fell” could return “tell” “fall” or “felt.”

Geopbyte: 1,024 brontobytes. See Byte.

Ghost: See Bit Stream Backup.

GIF (Graphics Interchange Format): A common file format for storing images first originated by CompuServe, an Internet service provider, in 1987. Limited to 256 colors.

Gigabyte (GB): 1,024 megabytes. See Byte.

Global Address List (GAL): A Microsoft Outlook directory of all Microsoft Exchange users and distribution lists to which messages can be addressed. The global address list may also contain public folder names. Entries from this list can be added to a user’s personal address book (PAB).

Global De-Deduplication: See Case De-Duplication.

Global Positioning System (GPS): A technology used to track the location of ground based objects using three or more orbiting satellites.

GMT Timestamp: Identification of a file using Greenwich Mean Time as the central time authentication method. See Normalization.

GPS Generated Timestamp: Timestamp identifying time as a function of its relationship to Greenwich Mean Time.

Graphical User Interface (GUI, pronounced “goeey”): An interface to a computer or device comprised of pictures and icons, rather than words and numbers by which users can interact with the device.

Grayscale: See Scale-to-Gray.

Groupware: Software designed to operate on a network and allow several people to work together on the same documents and files.

Hacker: Someone who breaks into a computer system in order to access, steal, change or destroy information.

Half Duplex: Transmission systems that can send and receive data between computers but not at the same time.

Handshake: A transmission that occurs at the beginning of a communication session between computers to establish the technical format of the communication.

Handwriting Recognition Software (HRS): Software that interprets handwriting into machine readable form.

Hard Drive: A storage device consisting of one or more magnetic media platters on which digital data can be written and erased. See Platter.

Harvesting: The process of retrieving or collecting Electronically Stored Information from any media; an e-discovery vendor or specialist “harvests” Electronically Stored Information from computer hard drives, file servers, CDs, backup tapes, portable devices, and other sources for processing and loading to storage media or a database management system.

Hash Coding: A mathematical algorithm that calculates a unique value for a given set of data, similar to a digital fingerprint, representing the binary content of the data to assist in subsequently ensuring that data has not been modified. Common hash algorithms include MD5 and SHA. See Data Verification, Digital Fingerprint, File Level Binary Comparison.

Head: Devices which ride very closely to the surface of the platter on a hard drive and allow information to be read from and written to the platter.

Header: Data placed at the beginning of a file or section of data that in part identifies the file and some of its attributes. A header can consist of multiple fields, each containing its own value. See Message Header

Hexadecimal: A number system with a base of 16. The digits are 0-9 and A-F, where F equals the decimal value of 15.

Hidden Files or Data: Files or data not readily visible to the user of a computer. Some operating system files are hidden to prevent inexperienced users from inadvertently deleting or changing these essential files. See Steganography.

Hierarchical Storage Management (HSM): Software that automatically migrates files from on-line to less expensive near-line storage, usually on the basis of the age or frequency of use of the files.

High Technology Crime Investigation Association (HTCIA): Computer forensics nonprofit association; resources include educational programs and Listservs.

Hold: See Legal Hold.

Horizontal De-Duplication: A way to identify Electronically Stored Information duplicated across multiple custodians or other production data sets normally by comparing hash algorithms to identify duplicates and then removing or suppressing those duplicates. See Case De-Duplication, De-Duplication.

Host: In a network, the central computer that controls the remote computers and holds the central databases.

Hub: A network device that connects multiple computers and peripherals together allowing them to share network connectivity. A central unit that repeats and/or amplifies data signals being sent across a network.

Hyperlink: A pointer in a hypertext document – usually appearing as an underlined or highlighted word or picture – upon selection sends a user to another location either within the current document or to another location accessible on the network or Internet.

HyperText: Text that includes hyperlinks or shortcuts to other documents or views, allowing the reader to easily jump from one view to a related view in a non-linear fashion.

HyperText Markup Language (HTML): Developed by CERN of Geneva, Switzerland; the most common programming language format used on the Internet. HTML+ adds support for multi-media. The tag-based ASCII language used to create pages on the World Wide Web uses tags to tell a Web browser to display text and images. HTML is a markup or “presentation” language, not a programming language. Programming code can be imbedded in an HTML page to make it interactive. See Java.

HyperText Transfer Protocol (HTTP): The underlying protocol used by the World Wide Web. HTTP defines how messages are formatted and transmitted, and what actions servers and browsers should take in response to various commands. For example, when you enter a URL in your browser, this actually sends an HTTP command to the Web server directing it to fetch and transmit the requested site. HTTPS adds a layer of encryption to the protocol to protect the information that is being transmitted and is often used by application service providers to protect the data being viewed over the Web.

Icon: In a GUI, a picture or drawing that is activated by clicking a mouse to command the computer program to perform a predefined series of actions.

Image: (1) To make an identical copy of a storage device, including empty sectors. Also known as creating a mirror image or mirroring the drive. See Bit Stream Backup, Forensic Copy, Mirror Image. (2) An electronic or digital picture of a document (e.g., TIFF, PDF, etc.). See Image Processing, Processing Data, Render Images.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

Image Copy or Imaged Copy: See Forensic Copy.

Image Enabling: A software function that creates links between existing applications and stored images.

Image File Format: See File Format, Format (noun).

Image Key: The name of an image and cross reference to the image's file in a document load file, often the Bates number of the page.

Image Processing: To convert data from its current/native format to a fixed image for the purposes of preserving the format of a document and facilitating the transfer between parties, typically with the addition of a Bates number to the face of each image. See Form of Production, Native Format, Processing Data, Render Images.

Image Processing Card (IPC): A board mounted in a computer, scanner or printer that facilitates the acquisition and display of images. The primary function of most IPCs is the rapid compression and decompression of image files.

Import: The process of bringing data into an environment or application that has been exported from another environment or application.

Inactive Record: Records related to closed, completed or concluded activities. Inactive records are no longer routinely referenced but must be retained in order to fulfill reporting requirements or for purposes of audit or analysis. Inactive records generally reside in a long-term storage format remaining accessible for purposes of business processing only with restrictions on alteration. In some business circumstances, inactive records may be reactivated.

Incremental Backup: A method of backing up data that backs up data that is new or has been changed from that last backup of any kind, be it a full backup or the last incremental backup.

Index/Coding Fields: Database fields used to categorize and organize records. Often user-defined, these fields can be used for searching for and retrieving records. See Coding.

Index: A searchable catalog of information created to maximize storage efficiency and allow for improved search. Also called catalog. See Full-Text Indexing.

Indexing: (1) The process of organizing data in a database to maximize storage efficiency and optimize searching; (2) Objective coding of documents to create a list similar to a table of contents. See Coding.

Information: For the purposes of this document, information is used to mean hard copy documents and ESI.

Information Governance: Information governance is the comprehensive, inter-disciplinary framework of policies, procedures, and controls used by mature organizations to maximize the value of an organization's information while minimizing associated risks by incorporating the requirements of: (1) e-discovery, (2) records & information management and (3) privacy/security into the process of making decisions about information.

Information Lifecycle Management (ILM): A phrase used to discuss the policies and procedures governing the management of data within an organization, from creation through destruction. See Electronic Document Management.

Information Systems (IS) or Information Technology (IT): Usually refers to the department of an entity which designs, maintains, and assists users with regard to the computer infrastructure.

Information Technology (IT) Infrastructure: The overall makeup of business-wide technology operations, including mainframe operations, standalone systems, email, networks (WAN and LAN), Internet access, customer databases, enterprise systems, application support, regardless of whether managed, utilized or provided locally, regionally, globally, etc., or whether performed or located internally or by outside providers (outsourced to vendors). The IT infrastructure also includes applicable standard practices and procedures, such as backup procedures, versioning, resource sharing, retention practices, system cleanup, and the like. See Enterprise Architecture.

Infrastructure as a Service (IaaS): A form of Cloud Computing whereby a third party service provider offers, on-demand, a part of its computer infrastructure remotely. Specific services may include servers, software or network equipment resources that can be provided on an as-needed basis without the purchase of the devices or the resources needed to support them. See Cloud Computing.

Input device: Any peripheral that allows a user to communicate with a computer by entering information or issuing commands (e.g., keyboard).

Instant Messaging (IM): A form of electronic communication involving immediate correspondence between two or more online users. Instant messages differ from email in their limited metadata and in that messages are often transitory and not stored past the messaging session.

Institute of Electrical and Electronic Engineers (IEEE): An international association that advocates the advancement of technology as it relates to electricity. IEEE sponsors meetings, publishes a number of journals, and establishes standards.

Integrated Drive Electronics (IDE): An engineering standard for interfacing computers and hard disks.

Intelligent Character Recognition (ICR): The conversion of scanned images (bar codes or patterns of bits) to computer recognizable codes (ASCII characters and files) by means of software/programs that define the rules of and algorithms for conversion, helpful for interpreting handwritten text. See Handwriting Recognition Software (HRS), OCR.

Inter-Partition Space: Unused sectors on a track located between the start of the partition and the partition boot record of a hard drive. This space is important because it is possible for a user to hide information here. See Partition, Track.

Interlaced: To refresh a display every other line once per refresh cycle. Since only half the information displayed is updated each cycle, interlaced displays are less expensive than non-interlaced. However, interlaced displays are subject to jitters. The human eye/brain can usually detect displayed images that are completely refreshed less than 30 times per second.

Interleave: To arrange data in a noncontiguous way to increase performance. When used to describe disk drives, it refers to the way sectors on a disk are organized. In one-to-one interleaving, the sectors are placed sequentially around each track. In two-to-one interleaving, sectors are staggered so that consecutively numbered sectors are separated by an intervening sector. The purpose of interleaving is to make the disk drive more efficient. The disk drive can access only one sector at a time, and the disk is constantly spinning beneath.

International Organization for Standardization (ISO): A worldwide federation of national standards bodies; www.iso.org.

International Telecommunication Union (ITU): An international organization under the UN, headquartered in Geneva, concerned with telecommunications that develops international data communications standards; known as CCITT prior to March 1, 1993. See <http://www.itu.int>.

Internet: A worldwide interconnected system of networks that all use the TCP/IP communications protocol and share a common address space. The Internet supports services such as email, the World Wide Web, file transfer (FTP), and Internet Relay Chat (IRC). Also known as “the net,” “the information superhighway,” and “cyberspace.”

Internet Protocol (IP): The principal communications protocol for data communications across the Internet.

Internet Protocol (IP) Address: A unique name that identifies the physical location of a server on a network, expressed by a numerical value (e.g., 128.24.62.1). See Transmission Control Protocol/Internet Protocol (TCP/IP).

Internet Publishing Software: Specialized software that allows materials to be published to the Internet. The term Internet Publishing is sometimes used to refer to the industry of online digital publication as a whole.

Internet Relay Chat (IRC): System allowing Internet users to chat in real time.

Internet Service Provider (ISP): A business that provides access to the Internet, usually for a fee.

Integrated Services Digital Network (ISDN): An all-digital network that can carry data, video, and voice.

Intranet: A secure private network that uses Internet-related technologies to provide services within an organization or defined infrastructure.

ISO 8859-1: Also called Latin-1. A standard character encoding of the Latin alphabet used for most Western European languages. ISO 8859-1 is considered a legacy encoding in relation to Unicode, yet it is nonetheless still in common use today. The ISO 8859-1 standard consists of 191 printable characters from the Latin script. It is essentially a superset of the ASCII character encoding and a subset of the Windows-1252 character encoding. See ASCII, Windows-1252.

ISO 9660 CD Format: The ISO format for creating CD-ROMs that can be read worldwide.

ISO 15489-1: The ISO standard addressing standardization of international best practices in records management.

Jailbreak: A process of bypassing security restrictions of an operating system to take full control of a device.

Java: A platform-independent, programming language for adding animation and other actions to websites.

Joint Photographic Experts Group (JPEG): A compression algorithm for still images that is commonly used on the Web.

Journal: A chronological record of data processing operations that may be used to reconstruct a previous or an updated version of a file. In database management systems, it is the record of all stored data items that have values changed as a result of processing and manipulation of the data.

Journaling: A function of electronic communication systems (such as Microsoft Exchange and Lotus Notes) that copies items that are sent and received into a second information store for retention or preservation. Because Journaling takes place at the information store (server) level when the items are sent or received, rather than at the mailbox (client) level, some message-related metadata, such as user foldering (what folder the item is stored in within the recipient's mailbox) and the status of the "read" flag, is not retained in the journaled copy. The Journaling function stores items in the system's native format, unlike email archiving solutions, that use proprietary storage formats designed to reduce the amount of storage space required. Journaling systems may also lack the sophisticated search and retrieval capabilities available with many email archiving solutions. See Email Archiving.

Jukebox: A mass storage device that holds optical disks and automatically loads them into a drive.

Jump Drive: See Flash Drive.

Kerning: Adjusting the spacing between two letters.

Key Drive: See Flash Drive.

Key Field: See Primary Key

Keyword: Any specified word, or combination of words, used in a search, with the intent of locating certain results.

Kilobyte (KB): A unit of 1,024 bytes. See Byte.

Kofax Board: The generic term for a series of image processing boards manufactured by Kofax Imaging Processing. These are used between the scanner and the computer and perform real-time image compression and decompression for faster image viewing, image enhancement, and corrections to the input to account for conditions such as document misalignment.

Landscape Mode: A page orientation or display such that the width exceeds the height (Horizontal).

Laser Disk: Same as an optical CD, except 12" in diameter.

Laser Printing: A printing process by which a beam of light hits an electrically charged drum and causes a discharge at that point. Toner is then applied, which sticks to the non-charged areas. Paper is pressed against the drum to form the image and is then heated to dry the toner.

Latency: The time it takes to read a disk (or jukebox), including the time to physically position the media under the read/write head, seek the correct address and transfer it.

Latent Data: Deleted files and other Electronically Stored Information that are inaccessible without specialized forensic tools and techniques. Until overwritten, these data reside on media such as a hard drive in unused space and other areas available for data storage. Also known as ambient data. See Residual Data.

Latent Semantic Indexing and Analysis: A method of processing data that identifies relationships between data sets by analyzing terms and term frequency. Common applications include grouping documents together based on the documents' concepts and meanings instead of by simple searching.

Latin-1: See ISO 8859-1.

Leading: The amount of space between lines of printed text.

Legacy Data, Legacy System: ESI which can only be accessed via software and/or hardware that has become obsolete or replaced. Legacy data may be costly to restore or reconstruct when required for investigation or litigation analysis or discovery.

Legal Hold: A legal hold is a communication issued as a result of current or reasonably anticipated litigation, audit, government investigation or other such matter that suspends the normal disposition or processing of records. Legal holds may encompass procedures affecting data that is accessible as well as data that is not reasonably accessible. The specific communication to business or IT organizations may also be called a hold, preservation order, suspension order, freeze notice, hold order, litigation hold, or hold notice. See, The Sedona Conference Commentary on Legal Holds, September 2010, available for download at <http://www.thesedonaconference.org/download-pub/470>.

Lempel-Ziv & Welch (LZW): A common, lossless compression standard for computer graphics, used for most TIFF files. Typical compression ratios are 4/1.

Level Coding: Used in bibliographical coding to facilitate different treatment, such as prioritization or more thorough extraction of data, for different categories of documents, such as by type or source. See Coding.

LFP: IPRO Tech Inc.'s image cross reference file; an ASCII delimited text file that cross references an image's Bates number to its location and file name.

Lifecycle: A record's lifecycle is the life span of a record from its creation or receipt to its final disposition. Usually described in three stages: (1) creation, (2) maintenance and use, and (3) archive to final disposition. See Information Lifecycle Management.

Linear and Non-Linear Review: Performed by humans. Linear review workflow begins at the beginning of a collection and addresses information in order until a full review of all information is complete. Non-linear review workflow is to prepare only certain portions for review, based either on the results of criteria, such as search terms, computer assisted review results or some other method, to isolate only information likely responsive. See Review.

Linear Tape-Open (LTO): A type of magnetic backup tape that can hold as much as 800 GB of data, or 1200 CDs depending on the data file format.

Link: See Hyperlink.

Liquid Crystal Display (LCD): Two polarizing transparent panels with a liquid crystal surface between; application of voltage to certain areas causes the crystal to turn dark, and a light source behind the panel transmits though crystals not darkened.

Litigation Hold: See Legal Hold.

Load File: A file that relates to a set of scanned images or electronically processed files that indicates where individual pages or files belong together as documents, to include attachments, and where each document begins and ends. A load file may also contain data relevant to the individual documents, such as selected metadata, coded data, and extracted text. Load files should be obtained and provided in prearranged or standardized formats to ensure transfer of accurate and usable images and data.

Cited: *Aguilar v. Immigration and Customs Enforcement Division of the United States Department of Homeland Security*, 255 F.R.D. 350, 353 (S.D.N.Y. 2008).

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

Cited: *E.E.O.C. v. SVT, LLC*, 2014 U.S. Dist. LEXIS 50114 (N.D. Ind. Apr. 10, 2014).

Cited: *Nat'l Day Laborer Org. Network v. United States Immigration & Customs Enforcement Agency*, 2011 U.S. Dist. LEXIS 11655 (S.D.N.Y. February 7, 2011).

Local Area Network (LAN): A group of computers at a single location (usually an office or home) that are connected by phone lines, coaxial cable or wireless transmission. See Network.

Log File: A text file created by an electronic device or application to record activity of a server, website, computer or software program.

Logical Entities: An abstraction of a real-world object or concept that is both independent and unique. Conceptually, a logical entity is a noun, and its relationships to other entities are verbs. In a relational database, a logical entity is represented as a table. Attributes of the entity are in columns of the table and instances of the entity are in rows of the table. Examples of logical entities are employees of a company, products in a store's catalog, and patients' medical histories.

Logical File Space: The actual amount of space occupied by a file on a hard drive. The amount of logical file space differs from the physical file space because when a file is created on a computer, a sufficient number of clusters (physical file space) are assigned to contain the file. If the file (logical file space) is not large enough to completely fill the assigned clusters (physical file space), then some unused space will exist within the physical file space.

Logical Unitization: See Unitization – Physical and Logical.

Logical Volume: An area on the hard drive that has been formatted for file storage. A hard drive may contain a single or multiple volumes.

Lossless Compression: A method of compressing an image file, bit by bit, that results in no loss of information either during compression or extraction.

Lossy Compression: A method of image compression whereby storage size of image is reduced by decreasing the resolution and color fidelity while maintaining minimum acceptable standard for general use. A lossy image is one where the image after compression is different from the original image due to lost information. The differences may or may not be noticeable, but a lossy conversion process does not retain all the original information. JPEG is an example of a lossy compression method.

Lotus Domino: An IBM server product providing enterprise-level email, collaboration capabilities and custom application platform; began life as Lotus Notes Server, the server component of Lotus Development Corporation's client-server messaging technology. Can be used as an application server for Lotus Notes applications and/or as a Web server. Has a built-in database system in the format of .nsf.

Magnetic/Optical Storage Media: The physical piece of material that receives data that has been recorded using a number of different magnetic recording processes. Examples include hard drives, backup tapes, CD-ROMs, DVD-ROMs, Jaz and Zip Drives.

Magnetic Disk Emulation (MDE): Software that makes a jukebox look and operate like a hard drive such that it will respond to all the input/output (I/O) commands ordinarily sent to a hard drive.

Magneto-Optical Drive: A drive that combines laser and magnetic technology to create high-capacity erasable storage.

Mail Application Programming Interface (MAPI): A Windows-based software standard that enables a program to send and receive email by connecting the program to selected email servers. See API.

Mailbox: A term used to describe all email associated with an individual email account, whether located physically together on one server, across a server array or stored in cloud based storage.

Make-Available Production: Process by which a generally large universe of potentially responsive documents is made available to a requestor; the requestor selects or tags desired documents, and the producing party produces only the selected documents. See Quick Peek.

Malware: Any type of malicious software program, typically installed illicitly, including viruses, Trojans, worms, key loggers, spyware, adware and others.

Management Information Systems (MIS): A phrase used to describe the resources, people and technology, used to manage the information of an organization.

MAPI Mail Near-Line: Documents stored on optical disks or compact disks that are housed in a jukebox or CD changer and can be retrieved without human intervention.

Marginalia: Handwritten notes on documents.

Master Boot Sector/Record: The sector on a hard drive that contains the computer code (boot strap loader) necessary for the computer to start up and the partition table describing the organization of the hard drive.

Master File Table (MFT): The primary record of file storage locations on a Microsoft Windows based computer employing NTFS filing systems.

Mastering: Making many copies of a disk from a single master disk.

MBOX: The format in which email is stored on traditional UNIX email systems.

Media: An object or device, such as a disk, tape or other device, on which data is stored.

Megabyte (MB): 1,024 kilobytes. See Byte.

Memory: Data storage in the form of chips, or the actual chips used to hold data; storage is used to describe memory that exists on tapes, disks, CDs, DVDs, flash drives and hard drives. See Random Access Memory (RAM), Read Only Memory (ROM).

Menu: A list of options, each of which performs a desired action such as choosing a command or applying a particular format to a part of a document.

Message-Digest Algorithm 5 (MD5): A hash algorithm used to give a numeric value to a digital file or piece of data. Commonly used in electronic discovery to find duplicates in a data collection. See Hash Coding.

Message Header: The text portion of an email that contains routing information of the email and may include author, recipient and server information, which tracks the path of the email from its origin server to its destination mailbox.

Message Unit: An email and any attachments associated with it.

Metadata: The generic term used to describe the structural information of a file that contains data about the file, as opposed to describing the content of a file. See System Generated Metadata and User Created Metadata. For a more thorough discussion, see *The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age (Second Edition)*, <https://thesedonaconference.org/download-pub/74>, and *The Sedona Conference Commentary on Ethics & Metadata*, <https://thesedonaconference.org/download-pub/3111>.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

Metadata Comparison: A comparison of specified metadata as the basis for de-duplication without regard to content. See De-Duplication.

Microfiche: Sheet microfilm (4" by 6") containing reduced images of 270 pages or more in a grid pattern.

Microprocessor: See CPU.

Microsoft-Disk Operating System (MS-DOS): Used in Windows-based personal computers as the control system prior to the introduction of 32-bit operating systems.

Microsoft Outlook: A personal information manager from Microsoft, part of the Microsoft Office suite. Although often used mainly as an email application, it also provides calendar, task and contact management, note taking, a journal and Web browsing. Can be used as a stand-alone application, or operate in conjunction with Microsoft Exchange Server to provide enhanced functions for multiple users in an organization, such as shared mailboxes and calendars, public folders and meeting time allocation.

Microsoft Outlook Express: A scaled down version of Microsoft Outlook.

MiFi: A portable wireless hub that allows users with the correct credentials to access the Internet.

Migrated Data: Electronically Stored Information that has been moved from one database or format to another.

Migration: Moving Electronically Stored Information from one computer application or platform to another; may require conversion to a different format.

Mirror Image: A bit by bit copy of any storage media. Often used to copy the configuration of one computer to another computer or when creating a preservation copy. See Forensic Copy and Image.

Cited: *White v. Graceland College Center for Professional Development & Lifelong Learning, Inc.*, 2009 WL 722056 at *6 (D. Kan. March 18, 2009).

Mirroring: The duplication of Electronically Stored Information for purposes of backup or to distribute Internet or network traffic among several servers with identical ESI. See Bit Stream Backup, Disk Mirroring, Image.

Modem (Modulator-Demodulator): A device that can encode digital information into an analog signal (modulates) or decode the received analog signal to extract the digital information (demodulate).

Mount or Mounting: The process of making off-line Electronically Stored Information available for online processing. For example, placing a magnetic tape in a drive and setting up the software to recognize or read that tape. The terms load and loading are often used in conjunction with, or synonymously with, mount and mounting (as in “mount and load a tape”). Load may also refer to the process of transferring Electronically Stored Information from mounted media to another media or to an online system.

MPEG-1, -2, -3 and -4: Different standards for full motion video to digital compression/decompression techniques advanced by the Moving Pictures Experts Group.

MSG: A common file format, in which emails can be saved, often associated with Microsoft Outlook email program, which preserves both the format and any associated attachment information.

Multimedia: The combined use of different media; integrated video, audio, text and data graphics in digital form.

Multimedia Messaging Service (MMS): A protocol of messaging that allows for the transmission of multimedia content such as pictures, video or sound over mobile networks. See Text Message.

Multisync: Analog video monitors that can receive a wide range of display resolutions, usually including TV (NTSC). Color analog monitors accept separate red, green & blue (RGB) signals.

National Institute of Standards and Technology (NIST): A federal technology agency that works with industry to develop and apply technology measurements and standards. See NIST List.

Native Format: Electronic documents have an associated file structure defined by the original creating application. This file structure is referred to as the native format of the document. Because viewing or searching documents in the native format may require the original application (for example, viewing a Microsoft Word document may require the Microsoft Word application), documents may be converted to a neutral format as part of the record acquisition or archive process. Static format (often called imaged format), such as TIFF or PDF, are designed to retain an image of the document as it would look viewed in the original creating application but do not allow metadata to be viewed or the document information to be manipulated unless agreed-upon metadata and extracted text are preserved. In the conversion to static format, some metadata can be processed, preserved and electronically associated with the static format file. However, with technology advancements, tools and applications are increasingly available to allow viewing and searching of documents in their native format while still preserving pertinent metadata. It should be noted that not all ESI may be conducive to production in either the Native Format or imaged format, and some other form of production may be necessary. Databases, for example, often present such issues. See Form of Production, Load File.

Cited: *Akanthos Capital Management, LLC v. CompuCredit Holdings Corp.*, 2014 WL 896743 (N.D. Ga. Mar. 7, 2014).

Cited: *Covad Communications Co. v. Revonet, Inc.*, 254 F.R.D. 147 (D.D.C. 2008).

Cited: *Palar v. Blackhawk Bancorporation Inc.*, 2013 WL 1704302 (C.D. Ill. Mar. 19, 2013)

Cited: *Race Tires America, Inc. v. Hoosier Racing Tire Corp.*, 2012 WL 887593 (3d Cir. March 16, 2012).

Native Format Review: Review of Electronically Stored Information in its native format using either a third party viewer application capable of rendering native files in close approximation to their original application or the actual original application in which the Electronically Stored Information was created. See Review.

Natural Language Search: A manner of searching that permits the use of plain language without special connectors or precise terminology, such as “Where can I find information on William Shakespeare?” as opposed to formulating a search statement (such as “information” and “William Shakespeare”). See Boolean Search.

Near Duplicates: (1) Two or more files that are similar to a certain percentage, for example, files that are 90% similar may be identified as near duplicates; used for review to locate similar documents and review all near duplicates at one time; (2) The longest email in an email conversation where the subparts are identified and suppressed in an email collection to reduce review volume.

Near-Line Data Storage: A term used to refer to a data storage system where data is not actively available to users, but is available through an automated system that enables the robotic retrieval of removable storage media or tapes. Data in near-line storage is often

stored on servers that do not have as high performance as active servers. Making near-line data available will not require human intervention (as opposed to off-line data which can only be made available through human actions).

Network: A group of two or more computers and other devices connected together (“networked”) for the exchange and sharing of resources. See Local-Area Network (LAN) and Wide-Area Network (WAN).

Network Operating System (NOS): See Operating System.

Network Operations Center (NOC): The location where a network or computer array is monitored and maintained.

Neural Network: Neural networks are made up of interconnected processing elements called units, which respond in parallel to a set of input signals given to each.

New Technology File System (NTFS): A high-performance and self-healing file system proprietary to Microsoft, used in Windows NT, Windows 2000, Windows XP and Windows Vista Operating Systems, that supports file-level security, compression and auditing. It also supports large volumes and powerful storage solutions such as Redundant Array of Inexpensive Disks (RAID). An important feature of NTFS is the ability to encrypt files and folders to protect sensitive data.

NIST List: A hash database of computer files developed by NIST to identify files that are system generated and generally accepted to have no substantive value in most instances.

Node: Any device connected to a network. PCs, servers and printers are all nodes on the network.

Non-Interlace: When each line of a video image is scanned separately. Older CRT computer monitors use non-interlaced video.

Normalization: The process of reformatting data so that it is stored in a standardized form, such as setting the date and time stamp of a specific volume of Electronically Stored Information to a specific zone, often GMT, to permit advanced processing of the ESI, such as de-duplication. See Coordinated Universal Time.

Notes Server: See Lotus Domino.

NSF: Lotus Notes container file (i.e., database.nsf); can be either an email database or the traditional type of fielded database. See Lotus Domino.

Object: In personal computing, an object is a representation of something that a user can work with to perform a task and can appear as text or an icon. In a high-level method of programming called object-oriented programming (OOP), an object is a freestanding block of code that defines the properties of something.

Object Linking and Embedding (OLE): A feature in Microsoft’s Windows that allows the linking of different files, or parts of files, together into one file without forfeiting any of the original file’s attributes or functionality. See Compound Document.

Objective Coding: See Coding.

OCR (Optical Character Recognition): A technology process that captures text from an image for the purpose of creating a parallel text file that can be associated with the image and searched in a database. OCR software evaluates scanned data for shapes it recognizes as letters or numerals. See Handwriting Recognition Software (HRS), Intelligent Character Recognition (ICR).

Cited (Glossary Third Ed.): *Country Vintner of North Carolina, LLC v. E. & J. Gallo Winery, Inc.*, 718 F.3d 249 (4th Cir. Apr. 29, 2013).

Official Record Owner: See Record Owner.

Offline Data: The storage of Electronically Stored Information outside the network in daily use (e.g., on backup tapes) that is only accessible through the offline storage system, not the network.

Offline Storage: Electronically Stored Information stored on removable disk (optical, compact, etc.) or magnetic tape and not accessible by the active software or server. Often used for making disaster-recovery copies of records for which retrieval is unlikely. Accessibility to offline media usually requires restoring the data back to the active server.

Online: Connected to a network or the Internet.

Online Review: The review of data on a computer, either locally on a network or via the Internet. See Review.

Online Storage: The storage of Electronically Stored Information as fully accessible information in daily use on the network or elsewhere.

Ontology: A collection of categories and their relationships to other categories and to words. An ontology is one of the methods used to find related documents when given a specific query.

Operating System (OS): Provides the software platform that directs the overall activity of a computer, network or system and on which all other software programs and applications run. In many ways, choice of an operating system will effect which applications can be run. Operating systems perform basic tasks, such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk and controlling peripheral devices such as disk drives and printers. For large systems, the operating system has even greater responsibilities and powers – becoming a traffic cop to make sure different programs and users running at the same time do not interfere with each other. The operating system is also responsible for security, ensuring that unauthorized users do not access the system. Examples of computer operating systems are UNIX, DOS, Microsoft Windows, LINUX, Mac OS and IBM z/OS. Examples of portable device operating systems are iOS, Android, Microsoft Windows and BlackBerry. Operating systems can be classified in a number of ways, including: multi-user (allows two or more users to run programs at the same time – some operating systems permit hundreds or even thousands of concurrent users); multiprocessing (supports running a program on more than one CPU); multitasking (allows more than one program to run concurrently); multithreading (allows different parts of a single program to run concurrently); and real time (instantly responds to input – general-purpose operating systems, such as DOS and UNIX, are not real-time).

Optical Disks: Computer media similar to a compact disk that cannot be rewritten. An optical drive uses a laser to read the ESI.

Optical Jukebox: See Jukebox.

OST: A Microsoft Outlook information store that is used to save folder information that can be accessed offline.

Cited: *White v. Graceland College Center for Professional Development & Lifelong Learning, Inc.*, 2009 WL 722056 at *5 (D. Kan. March 18, 2009).

Outlook: See Microsoft Outlook.

Over-inclusive: When referring to data sets returned by some method of query, search, filter or cull, results that are returned overly broad.

Overwrite: To record or copy new data over existing data, as in when a file or directory is updated.

Packet: A unit of data sent across a network that may contain identity and routing information. When a large block of data is to be sent over a network, it is broken up into several packets, sent and then reassembled at the other end. The exact layout of an individual packet is determined by the protocol being used.

Page File/Paging File: A method to temporarily store data outside of the main memory but quickly retrievable. This information is left in the swap file after the programs are terminated and may be retrieved using forensic techniques. See Swap File.

Parallel Port: See Port.

Parent: See Document.

Parsing: In electronic discovery, the process by which a file is broken apart into its individual components for indexing, processing or to prepare for loading into a review database.

Partition: An individual section of computer storage media such as a hard drive. For example, a single hard drive may be divided into several partitions in order that each partition can be managed separately for security or maintenance purposes. When a hard drive is divided into partitions, each partition is designated by a separate drive letter, i.e., C, D, etc.

Partition Table: Indicates each logical volume contained on a disk and its location.

Partition Waste Space: After the boot sector of each volume or partition is written to a track, it is customary for the system to skip the rest of that track and begin the actual useable area of the volume on the next track. This results in unused or wasted space on that track where information can be hidden. This wasted space can only be viewed with a low level disk viewer. However, forensic techniques can be used to search these wasted space areas for hidden information.

Password: A text or alphanumeric string that is used to authenticate a specific user's access to a secure program, network or part of a network.

Path: (1) The hierarchical description of where a directory, folder or file is located on a computer or network; (2) A transmission channel, the path between two nodes of a network that a data communication follows, and the physical cabling that connects the nodes on a network.

Pattern Matching: A generic term that describes any process that compares one file's content with another file's content.

Pattern Recognition: Technology that searches Electronically Stored Information for like patterns and flags, and extracts the pertinent data, usually utilizing an algorithm. For instance, in looking for addresses, alpha characters followed by a comma and a space, followed by two capital alpha characters, followed by a space, followed by five or more digits, are usually the city, state and zip code. By programming the application to look for a pattern, the information can be electronically identified, extracted, or otherwise utilized or manipulated.

Peer-to-Peer or P2P: A form of network organization that uses portions of each user's resources, like storage space or processing power, for use by others on the network. Notorious examples include the storage sharing of Napster or BitTorrent.

Peripheral: Any accessory device attached to a computer, such as a disk drive, printer, modem or joystick.

Peripheral Component Interconnect or Interface (PCI): A high-speed interconnect local bus used to support multimedia devices.

Personal Address Book (PAB): A file type to describe a Microsoft Outlook list of contacts created and maintained by an individual user for personal use.

Personal Computer (PC): Computer based on a microprocessor and designed to be used by one person at a time.

Personal Computer Memory Card International Association (PCMCIA): Plug-in cards for computers (usually portables) that extend the storage and/or functionality.

Personal Data (as used in the EU Data Protection Act): Data which relate to a natural person who can be identified from the data, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his or her physical, physiological, mental, economic, cultural or social identity. Also referred to as PII (Personally Identifiable Information).

Personal Digital Assistant (PDA): A portable device used to perform communication and organizational tasks.

Personal Filing Cabinet (PFC): The AOL proprietary email storage container file used for the local storage of emails, contacts, calendar events and other personal information.

Petabyte (PB): 1,024 terabytes (approximately one million gigabytes). See Byte.

Physical Disk: An actual piece of computer media, such as the hard disk or drive, floppy disks, CD-ROM disks, zip drive, etc.

Physical File Storage: When a file is created on a computer, a sufficient number of clusters are assigned to contain the file. If the file is not large enough to completely fill the assigned clusters then some unused space will exist within the physical file space. This unused space is referred to as file slack and can contain unused space, previously deleted/overwritten files or fragments thereof.

Physical Unitization: See Unitization – Physical and Logical.

Picture Element: The smallest addressable unit on a display screen. The higher the resolution (the more rows and columns), the more information can be displayed.

Ping: Executable command, used as a test for checking network connectivity.

Pitch: Characters (or dots) per inch, measured horizontally.

Pixel: A single unit of a raster image that allows a picture to be displayed on an electronic screen or computer monitor.

Plaintext or Plain Text: The least formatted and therefore most portable form of text for computerized documents.

Plasma Display: A type of flat panel display commonly use for large televisions in which many tiny cells are located between two panels of glass holding an inert mixture of gases that are then electronically charged to produce light.

Platform as a Service (PaaS): A form of Cloud computing which describes the outsourcing of the computer platform upon which development and other workflows can be performed without the costs of hardware, software and personnel. See Cloud Computing.

Platter: One of several components that make up a computer hard drive. Platters are thin, rapidly rotating disks that have a set of read/write heads on both sides of each platter. Each platter is divided into a series of concentric rings called tracks. Each track is further divided into sections called sectors, and each sector is sub-divided into bytes.

Plug and Play (PNP): A method by which new hardware may be detected, configured, and used by existing systems upon connection with little or no user intervention.

Pointer: An index entry in the directory of a disk (or other storage medium) that identifies the space on the disk in which an electronic document or piece of electronic data resides, thereby preventing that space from being overwritten by other data. In most cases, when an electronic document is deleted, the pointer is deleted, allowing the document to be overwritten, but the document is not actually erased until overwritten.

Port: An interface between a computer and other computers or devices. Can be divided into two primary groups based on signal transfer: Serial ports send and receive one bit at a time via a single wire pair, while parallel ports send multiple bits at the same time over several sets of wires. See Universal Serial Bus (USB) Port. Software ports are virtual data connections used by programs to exchange data directly instead of going through a file or other temporary storage locations; the most common types are Transmission Control Protocol/Internet Protocol (TCP/IP) and User Datagram Protocol (UDP).

Portable Document Format (PDF): A file format technology developed by Adobe Systems to facilitate the exchange of documents between platforms regardless of originating application by preserving the format and content.

Cited: *Country Vintner of North Carolina, LLC v. E. & J. Gallo Winery, Inc.*, 718 F.3d 249 (4th Cir. Apr. 29, 2013).

Cited: *Saliga v. Chemtura Corp.*, 2013 WL 6182227 (D. Conn. Nov. 25, 2013).

Portable Volumes: A feature that facilitates the moving of large volumes of documents without requiring copying multiple files. Portable volumes enable individual CDs to be easily regrouped, detached and reattached to different databases for a broader information exchange.

Portrait Mode: A page orientation or display such that the height exceeds the width (vertical).

Precision: When describing search results, precision is the number of documents retrieved from a search divided by the total number of documents returned. For example, in a search for documents relevant to a document request, it is the percentage of documents returned from a search that are actually relevant to the request. See The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, available for download at <https://thesedonaconference.org/download-pub/3669>.

Predictive Coding/Ranking: See Technology-Assisted Review.

Preservation: The process of retaining documents and ESI, including document metadata, for legal purposes and includes suspension of normal document destruction policies and procedures. See Spoliation.

Preservation Notice, Preservation Order: See Legal Hold.

Primary Key: A unique value stored in a field or fields of a database record that is used to identify the record and, in a relational database, to link multiple tables together.

Print On Demand (POD): A term referring to document images stored in electronic format and available to be quickly printed.

Printout: Printed data, also known as hard copy.

Private Network: A network that is connected to the Internet but is isolated from the Internet with security measures allowing use of the network only by persons within the private network.

Privilege Data Set: The universe of documents identified as responsive and/or relevant but withheld from production on the grounds of legal privilege, a log of which is usually required to notify of withheld documents and the grounds on which they were withheld (e.g., work product, attorney-client privilege).

Process/processing (as used in the EU Data Protection Act): Any operation or set of operations which is performed upon Personal Data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.

Processing Data: The automated ingestion of Electronically Stored Information into a program for the purpose of extracting metadata and text; and in some cases, the creation of a static image of the source Electronically Stored Information files according to a predetermined set of specifications, in anticipation of loading to a database. Specifications can include the de-duplication of ESI, or filtering based on metadata contents such as date or email domain and specific metadata fields to be included in the final product.

Production: The process of delivering to another party, or making available for that party's review, documents and/or Electronically Stored Information deemed responsive to a discovery request.

Production Data Set: The universe of documents and/or Electronically Stored Information identified as responsive to document requests and not withheld on the grounds of privilege.

Production Number: Often referred to as the Bates number. A sequential number assigned to every page of a production for fixed image production formats, or to every file in a native file production, used for tracking and reference purposes. Often used in conjunction with a suffix or prefix to identify the producing party, the litigation, or other relevant information. See Bates Number, Beginning Document Number.

Program: See Application and Software.

Properties: File level metadata describing attributes of the physical file, i.e., size, creation data and author. See Metadata.

Protocol: Defines a common series of rules, signals and conventions that allow different kinds of computers and applications to communicate over a network. One of the most common protocols for networks is called Transmission Control Protocol/Internet Protocol (TCP/IP).

Protodigital: Primitive or first-generation digital. Applied as an adjective to systems, software, documents, or ways of thinking. The term was first used in music to refer to early computer synthesizers that attempted to mimic the sound of traditional musical instruments and to early jazz compositions written on computers with that instrumentation in mind. In electronic discovery, this term is most often applied to systems or ways of thinking that – on the surface – appear to embrace digital technology, but attempt to equate Electronically Stored Information to paper records, ignoring the unique attributes of ESI. When someone says, “What’s the big deal with e-discovery? Sure we have a lot of email. You just print it all out and produce it like you used to”; that is an example of protodigital thinking. When someone says, “We embrace electronic discovery. We scan everything to .PDF before we produce it”; that person is engaged in protodigital thinking – attempting to fit Electronically Stored Information into the paper discovery paradigm.

Proximity Search: A search syntax written to find two or more words within a specified distance from each other.

PST: A Microsoft Outlook email storage file containing archived email messages in a compressed format.

*Cited: White v. Graceland College Center for Professional Development & Lifelong Learning, Inc., 2009 WL 722056 at *5 (D. Kan. March 18, 2009).*

Public Key Infrastructure (PKI) Digital Signature: A system, including hardware, software and policies, designed to manage digital certificates and match those certificates to specific users so that data can be validated as authentic. See Certificate, Digital Certificate and Digital Signature.

Public Network: A network that is part of the public Internet.

Quality Control (QC): Steps taken to ensure that results of a given task, product or service are of sufficiently high quality; the operational techniques and activities that are used to fulfill requirements for quality. In document handling and management processes, this includes image quality (resolution, skew, speckle, legibility, etc.), and data quality (correct information in appropriate fields, validated data for dates, addresses, names/issues lists, etc.).

Quarter Inch Cartridge (QIC): Digital recording tape, 2000 feet long, with an uncompressed capacity of 5 GB.

Query: An electronic search request for specific information from a database or other ESI.

Query By Image Content (QBIC): An IBM search system for stored images that allows the user to sketch an image and then search the image files to find those which most closely match. The user can specify color and texture – such as sandy beaches or clouds.

Queue: A sequence of items such as packets or print jobs waiting to be processed. For example, a print queue holds files that are waiting to be printed.

Quick Peek: An initial production whereby documents and/or Electronically Stored Information are made available for review or inspection before being reviewed for responsiveness, relevance, privilege, confidentiality or privacy. See Make-Available Production.

Quick Response (QR) Code: A small, square matrix pattern that can be read by an optical scanner or mobile phone camera; can store thousands of alphanumeric characters and may be affixed to business cards, advertising, product parts or other objects in order to convey information, commonly an internet URL.

Random Access Memory (RAM): Hardware inside a computer that retains memory on a short-term basis and stores information while the computer is in use. It is the working memory of the computer into which the operating system, startup applications and drivers are loaded when a computer is turned on, or where a program subsequently started up is loaded, and where thereafter, these applications are executed. RAM can be read or written in any section with one instruction sequence. It helps to have more of this working space installed when running advanced operating systems and applications to increase operating efficiency. RAM content is erased each time a computer is turned off. See Dynamic Random Access Memory (DRAM).

Raster/Rasterized (Raster or Bitmap Drawing): A method of representing an image with a grid (or map) of dots. Typical raster file formats are GIF, JPEG, TIFF, PCX, BMP, etc. and they typically have jagged edges.

Read Only Memory (ROM): Random memory that can be read but not written or changed. Also, hardware, usually a chip, within a computer containing programming necessary for starting the computer and essential system programs that neither the user nor

the computer can alter or erase. Information in the computer's ROM is permanently maintained even when the computer is turned off.

Recall: When describing search results, recall is number of documents retrieved from a search divided by all of the responsive documents in a collection. For example, in a search for documents relevant to a document request, it is the percentage of documents returned from a search compared against all documents that should have been returned and exist in the data set. See The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, available for download at <https://thesedonaconference.org/download-pub/3669>.

Record: (1) Information, regardless of medium or format that has value to an organization.
(2) A single row of information or subset of data elements in a database.

Record Custodian: An individual responsible for the physical storage of records throughout their retention period. In the context of electronic records, custodianship may not be a direct part of the records management function in all organizations. For example, some organizations may place this responsibility within their Information Technology Department, or they may assign responsibility for retaining and preserving records with individual employees. See Record Owner.

Record Lifecycle: The time period from which a record is created until it is disposed. See Information Lifecycle Management.

Record Owner: The physical custodian or subject matter expert on the contents of the record and responsible for the lifecycle management of the record. This may be, but is not necessarily, the author of the record. See Record Custodian.

Record Series: A description of a particular set of records within a file plan. Each category has retention and disposition data associated with it, applied to all record folders and records within the category. See DoD 5015.

Record Submitter: The person who enters a record in an application or system. This may be, but is not necessarily, the author or the record owner.

Records Archive: See Repository for Electronic Records.

Records Hold: See Legal Hold.

Records Management: The planning, controlling, directing, organizing, training, promoting and other managerial activities involving the life-cycle of information, including creation, maintenance (use, storage, retrieval) and disposition, regardless of media. See Information Lifecycle Management.

Records Manager: The person responsible for the implementation of a records management program in keeping with the policies and procedures that govern that program, including the identification, classification, handling and disposition of the organization's records throughout their retention life-cycle. The physical storage and protection of records may be a component of this individual's functions, but it may also be delegated to someone else. See Record Custodian.

Records Retention Period: The length of time a given record series must be kept, expressed as either a time period (e.g., four years), an event or action (e.g., audit), or a combination (e.g., six months after audit).

Records Retention Schedule: A plan for the management of records listing types of records and how long they should be kept; the purpose is to provide continuing authority to dispose of or transfer records to historical archives. See Information Lifecycle Management.

Records Store: See Repository for Electronic Records.

Recover, Recovery: See Restore.

Redaction: A portion of an image or document is intentionally obscured or removed to prevent disclosure of a specific portion. Done to protect privileged or irrelevant portions, including highly confidential, sensitive or proprietary information.

Redundant Array of Independent Disks (RAID): A method of storing data on servers that usually combines multiple hard drives into one logical unit, thereby increasing capacity, reliability and backup capability. RAID systems may vary in levels of redundancy, with no redundancy being a single, non-mirrored disk as level 0, two disks that mirror each other as level 1, on up, with level 5 being one of the most common. RAID systems are more complicated to restore and copy.

Refresh Rate: The number of times per second a computer display (such as on a CRT or TV) is updated.

Region (of an image): An area of an image file that is selected for specialized processing. Also called a zone.

Registration: (1) In document coding, it is the process of lining up an image of a form to determine the location of specific fields. See Coding; (2) entering pages into a scanner such that they are correctly read.

Relative Path: The electronic path on a network or computer to an individual file from a common point on the network.

Remote Access: The ability to access and use digital information from a location off-site from where the information is physically located; e.g., to use a computer, modem and some remote access software to connect to a network from a distant location.

Render Images: To take a native format electronic file and convert it to an image that appears as if the original format file were printed to paper. See Image Processing.

Replication: See Disk Mirroring.

Report: Formatted output of a system providing specific information.

Repository for Electronic Records: A direct access device on which the electronic records and associated metadata are stored. Sometimes called a records store or records archive.

Residual Data: Sometimes also referred to as Ambient Data; data that is not active on a computer system as the result of being deleted or moved to another location and is unintentionally left behind. Residual data includes: (1) data found on media free space; (2)

data found in file slack space; and (3) data within files that has functionally been deleted in that it is not visible using the application with which the file was created, without use of undelete or special data recovery techniques. May contain copies of deleted files, Internet files and file fragments. See Latent Data.

Resolution: Refers to the sharpness and clarity of an image. The term is most often used to describe monitors, printers and graphic images.

Restore: To transfer data from a backup medium (such as tapes) to an active system, often for the purpose of recovery from a problem, failure or disaster. Restoration of archival media is the transfer of data from an archival store to an active system for the purposes of processing (such as query, analysis, extraction or disposition of that data). Archival restoration of systems may require not only data restoration but also replication of the original hardware and software operating environment. Restoration of systems is often called recovery.

Retention Schedule: See Records Retention Schedule.

Reverse Engineering: The process of analyzing a system or piece of software to identify how it was created in order to recreate it in a new or different form. Reverse engineering is usually undertaken in order to redesign the system for better maintainability or to produce a copy of a system without utilizing the design from which it was originally produced. For example, one might take the executable code of a computer program, run it to study how it behaved with different input, and then attempt to write a program that behaved the same or better.

Review: The process of reading or otherwise analyzing documents to make a determination as to the document's applicability to some objective or subjective standard. Often used to describe the examination of documents in a legal context for their responsiveness or relevance to specific issues in a matter. See Native Format Review, Online Review.

Rewritable Technology: Storage devices where the data may be written more than once – typically hard drives, floppies and optical disks.

RFC822: Standard that specifies a syntax for text messages that are sent between one or more computer users, within the framework of email.

Rich Text Format (RTF): A standard text file format that preserves minimal stylistic formatting of document files for ease in exchange between various parties with different software.

Rip: To extract Electronically Stored Information from container files, for example to unbundle email collections into individual emails, during the e-discovery process while preserving metadata, authenticity and ownership. Also used to describe the extraction or copying of data to or from an external storage device.

RIM: Records and information management. [Also the acronym of the company that developed and sells BlackBerry devices, Research In Motion.]

RLE (Run Length Encoded): Compressed image format; supports only 256 colors; most effective on images with large areas of black or white.

Root Directory: The top level in a hierarchical file system. For example on a PC, the root directory of the hard drive, usually C:, contains all the second-level subdirectories on that drive.

Router: A device that forwards data packets along networks. A router is connected to at least two networks, commonly two LANs or WANs or a LAN and its ISPs network. Routers are located at gateways, the places where two or more networks connect. See Wireless Router.

SaaS (Software as a Service): Software application delivery model where a software vendor develops a Web-native software application and hosts and operates (either independently or through a third-party) the application for use by its customers over the Internet. Customers pay, not for owning the software itself, but for using it. See Application Service Provider (ASP), Cloud Computing.

Sampling: Sampling usually refers to the process of testing a database or a large volume of Electronically Stored Information for the existence or frequency of relevant information. It can be a useful technique in addressing a number of issues relating to litigation, including decisions about what repositories of data are appropriate to search in a particular litigation, and determinations of the validity and effectiveness of searches or other data extraction procedures.

SAN (Storage Area Network): A high-speed sub-network of shared storage devices. A storage device is a machine that contains nothing but a disk or disks for storing data. A SAN's architecture works in a way that makes all storage devices available to all servers on a LAN or WAN. As more storage devices are added to a SAN, they too will be accessible from any server in the larger network. The server merely acts as a pathway between the end user and the stored data. Because stored data does not reside directly on any of a network's servers, server power is utilized for business applications and network capacity is released to the end user. See Network.

SAS-70: (Statement on Auditing Standards No. 70, Service Organizations): An auditing standard developed by the American Institute of Certified Public Accountants (AICPA), which includes an examination of an entity's controls over information technology, security and related processes. There are two types of examinations: Type I examines the policies and procedures in place for their effectiveness to the stated objective; Type II reports on how the systems were actually used during the period of review. The SAS-70 Type II assessment is often used by hosting vendors and storage co-locations as a testament to their internal controls.

Scalability: The capacity of a system to expand without requiring major reconfiguration or re-entry of data. For example, multiple servers or additional storage can be easily added.

Scale-to-Gray: An option to display a black and white image file in an enhanced mode, making it easier to view. A scale-to-gray display uses gray shading to fill in gaps or jumps (known as aliasing) that occur when displaying an image file on a computer screen. Also known as grayscale.

Schema: A set of rules or conceptual model for data structure and content, such as a description of the data content and relationships in a database.

Scroll Bar: The bar on the side or bottom of a window that allows the user to scroll up and down through the window's contents. Scroll bars have scroll arrows at both ends and a scroll box, all of which can be used to scroll around the window.

Search: See Bayesian Search, Boolean Search, Compliance Search, Concept Search, Contextual Search, Full-Text Search, Fuzzy Search, Index, Keyword, Pattern Recognition, Proximity Search, Query By Image Content (QBIC), Sampling, Search Engine and Search Syntax.

Search Engine: A program that enables a search for keywords or phrases, such as on Web pages throughout the World Wide Web, e.g., Google, Bing, etc.

Search Syntax: The grammatical formatting of a search string, which is particular to the search program. Includes formatting for proximity searches, phrase searches or any other options that are supported by the search program.

Sector: A sector is normally the smallest individually addressable unit of information stored on a hard drive platter and usually holds 512 bytes of information. Sectors are numbered sequentially starting with 1 on each individual track. Thus, Track 0, Sector 1 and Track 5, Sector 1 refer to different sectors on the same hard drive. The first PC hard disks typically held 17 sectors per track.

Secure Hash Algorithm (SHA-1 and SHA-2): used for computing a condensed representation of a message or a data file specified by FIPS PUB 180. See Hash Coding.

Serial Line Internet Protocol (SLIP): A connection to the Internet in which the interface software runs in the local computer, rather than the Internet's.

Serial Port: See Port.

Server: Any central computer on a network that contains Electronically Stored Information or applications shared by multiple users of the network on their client computers; servers provide information to client machines. For example, there are Web servers that send out Web pages, mail servers that deliver email, list servers that administer mailing lists, FTP servers that hold FTP sites and deliver Electronically Stored Information to requesting users, and name servers that provide information about Internet host names. See File Server.

Server Farm: A cluster of servers.

Service-Level Agreement: A contract that defines the technical support or business parameters that a service provider or outsourcing firm will provide its clients. The agreement typically spells out measures for performance and consequences for failure.

Session: A lasting connection, usually involving the exchange of many packets between a user or host and a server, typically implemented as a layer in a network protocol, such as Telnet or File Transfer Protocol (FTP).

SGML/HyTime: A multimedia extension to SGML, sponsored by DoD.

Short Message Service (SMS): The most common data application for text messaging communication, allows users to send text messages to phones and other mobile communication devices. See Text Message.

Signature: See Certificate.

Simplex: One-sided page(s).

Single, In-Line Memory Module (SIMM): A mechanical package (with “legs”) used to attach memory chips to printed circuit boards.

Single Instance Storage: The method of de-duplication that is undertaken on a storage device to maximize space by eliminating multiple copies of a single file by retaining only one copy. This system of storage can occur either on a file level, or on a field level, where individual components of files are disassembled so that only unique parts are retained across an entire population and the reassembly of the original files is managed upon demand.

Siri: A personal assistant application for the Apple iPhone mobile device that allows users to search for the phone applications and Web services through voice recognition and natural language processing software.

Skewed: Tilted images. See De-skewing.

Slack/Slack Space: The unused space that exists on a hard drive when the logical file space is less than the physical file space. Also known as file slack. A form of residual data, the amount of on-disk file space from the end of the logical record information to the end of the physical disk record. Slack space can contain information soft-deleted from the record, information from prior records stored at the same physical location as current records, metadata fragments and other information useful for forensic analysis of computer systems. See Cluster Bitmap and Cluster (File).

Small Computer System Interface (SCSI, pronounced “skuzzy”): A common, industry standard connection type between computers and peripherals, such as hard disks, CD-ROM drives and scanners. SCSI allows for up to 7 devices to be attached in a chain via cables. SDLT (Super DLT): A type of backup tape that can hold up to 300 GB or 450 CDs, depending on the data file format. See Digital Linear Tape (DLT).

Smart Card: A credit card size device that contains a microprocessor, memory and a battery.

SMTP (Simple Mail Transfer Protocol): The protocol widely implemented on the Internet for exchanging email messages.

Snapshot: See Bit Stream Backup.

Social Network: A group of people that use the Internet to share and communicate, either professionally or personally, in a public setting typically based on a specific theme or interest. For example, Facebook is a popular social network that allows people to connect to friends and acquaintances anywhere in the world in order to share personal updates, pictures and experiences, and is used by entities as a public-facing presence.

Social Media: Internet applications which permit individuals or organizations to interactively share and communicate.

Software: Any set of coded instructions (programs) stored on computer-readable media that control what a computer does or can do. Includes operating systems and software applications.

Software application: See Application, Software.

Speckle: Imperfections in an image as a result of scanning paper documents that do not appear on the original. See De-speckling.

Spoliation: Spoliation is the destruction of records or properties, such as metadata, that may be relevant to ongoing or anticipated litigation, government investigation or audit. Courts differ in their interpretation of the level of intent required before sanctions may be warranted.

Cited: *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 269 F.R.D. 497(D. Md., 2010).

Cited: *Rimkus Consulting Group, Inc. v. Cammarata*, 688 F. Supp. 2d 598 (S.D. Tex., 2010).

Cited: *Quantlab Technologies Ltd. (BGI) v. Godlevsky*, 2014 WL 651944 (S.D. Tex. Feb. 19, 2014).

Spyware: A data collection program that secretly gathers information about the user and relays it to advertisers or other interested parties. Adware usually displays banners or unwanted pop-up windows, but often includes spyware as well. See Malware.

Stand-Alone Computer: A personal computer that is not connected to any other computer or network.

Standard Generalized Markup Language (SGML): An informal industry standard for open systems document management that specifies the data encoding of a document's format and content. Has been virtually replaced by Extensible Markup Language (XML).

Standard Parallel Port (SPP): See Port.

Steganography: The hiding of information within a more obvious kind of communication. Although not widely used, digital steganography involves the hiding of data inside a sound or image file. Steganalysis is the process of detecting steganography by looking at variances between bit patterns and unusually large file sizes.

Storage Device: A device capable of storing ESI.

Storage Media: See Magnetic/Optical Storage Media.

Streaming Indexing: Real-time or near real-time indexing of data as it being moved from one storage medium to another.

Structured Data: Data stored in a structured format, such as databases or data sets according to specific form and content rules as defined by each field of the database. Contrast to Unstructured Data.

Structured Query Language (SQL): A database computer language used to manage the data in relational databases. A standard fourth generation programming language (4GL – a programming language that is closer to natural language and easier to work with than a high-level language).

Subjective Coding: Recording the judgments of a reviewer as to a document's relevancy, privilege or importance with regard to factual or legal issues in a legal matter. See Coding.

Suspension Notice or Suspension Order: See Legal Hold.

Swap File: A file used to temporarily store code and data for programs that are currently running. This information is left in the swap file after the programs are terminated, and may be retrieved using forensic techniques. Also referred to as a page file or paging file.

System: (1) A collection of people, machines and methods organized to perform specific functions; (2) An integrated whole composed of diverse, interacting, specialized structures and sub-functions; and/or (3) A group of sub-systems united by some interaction or interdependence, performing many duties, but functioning as a single unit.

System Administrator (sysadmin, or sysop): The person responsible for and/or in charge of keeping a network or enterprise resource, such as a large database, operational.

System Files: Files allowing computer systems to run; non-user-created files.

System-Generated Metadata: Information about a file that is created and applied to a file by a computer process or application. Information could include the data a file was saved, printed or edited, and can include where a file was stored and how many times it has been edited. See Metadata.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

T1: A high speed, high bandwidth leased line connection to the Internet. T1 connections deliver information at 1.544 megabits per second.

T3: A high speed, high bandwidth leased line connection to the Internet. T3 connections deliver information at 44.746 megabits per second.

Tape Drive: A hardware device used to store or backup Electronically Stored Information on a magnetic tape. Tape drives are sometimes used to backup large quantities of Electronically Stored Information due to their large capacity and cheap cost relative to other storage options.

Taxonomy: The science of categorization, or classification, of things based on a predetermined system. In reference to Websites and portals, a site's taxonomy is the way it organizes its Electronically Stored Information into categories and subcategories, sometimes displayed in a site map. Used in information retrieval to find documents related to a query by identifying other documents in the same category.

Technology-Assisted Review (TAR)*: A process for prioritizing or coding a collection of Electronically Stored Information using a computerized system that harnesses human judgments of subject matter expert(s) on a smaller set of documents and then extrapolates those judgments to the remaining documents in the collection. Some TAR methods use algorithms that determine how similar (or dissimilar) each of the remaining documents is to those coded as relevant (or non-relevant, respectively) by the subject matter experts(s), while other TAR methods derive systematic rules that emulate the expert(s) decision-making process. TAR systems generally incorporate statistical models and/or sampling techniques to guide the process and to measure overall system effectiveness.

* Maura R. Grossman & Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review with Forward by John M. Facciola*, U.S. magistrate Judge, 2013 Fed. Cts. L. Rev. 7 (January 2013), <http://www.fclr.org/fclr/articles/html/2010/grossman.pdf>.

Telnet (Telecommunications Network): A protocol for logging onto remote computers from anywhere on the Internet.

Template: Sets of index fields for documents, providing framework for preparation.

Temporary (Temp) File: Contemporaneous files created by applications and stored on a computer for temporary use only; created to enable the processor of the computer to quickly pull back and assemble data for currently active files.

Terabyte: 1,024 gigabytes (approximately one trillion bytes). See Byte.

Text Delimited File: A common format for structured data exchange whereby a text file contains fielded data where the fields are separated by a specific ASCII character and also usually contain a header line that defines the fields contained in the file. See Field Separator or Field Delimiter.

Text Message: A written message typically restricted to 160 characters in length that is sent among users with mobile devices. The messages can be sent via the Short Messaging Service (SMS), as well as images, video and other multimedia using the Multimedia Messaging Service (MMS).

Text Mining: The application of data mining (knowledge discovery in databases) to unstructured textual data. Text mining usually involves structuring the input text (often parsing, along with application of some derived linguistic features and removal of others, and ultimate insertion into a database), deriving patterns within the data, and evaluating and interpreting the output, providing such ranking results as relevance, novelty, and interestingness. Also referred to as “Text Data Mining.” See Data Mining.

Text Retrieval Conference (TREC): An ongoing series of workshops co-sponsored by NIST and the U. S. Department of Defense.

TGA: Targa format. A scanned format – widely used for color-scanned materials (24-bit) as well as by various paint and desktop publishing packages.

Thin Client: A computer or software program which relies on a central server for processing and application resources, and Electronically Stored Information storage in a central area instead of locally; used mainly for output and input of user information or commands. See Client.

Thread: A series of technologically related communications, usually on a particular topic. Threads can be a series of bulletin board messages (for example, when someone posts a question and others reply with answers or additional queries on the same topic). A thread can also apply to emails or chats, where multiple conversation threads may exist simultaneously. See Email String.

Thumb Drive: See Flash Drive.

Thumbnail: A miniature representation of a page or item for quick overviews to provide a general idea of the structure, content and appearance of a document. A thumbnail program may be a standalone or part of a desktop publishing or graphics program. Thumbnails provide a convenient way to browse through multiple images before retrieving the one needed. Programs often allow clicking on the thumbnail to retrieve it.

TIFF (Tagged Image File Format): A widely used and supported graphic file formats for storing bit-mapped images, with many different compression formats and resolutions. File name has .TIF extension. Can be black and white, gray-scaled or color. Images are stored in tagged fields, and programs use the tags to accept or ignore fields, depending on the application. The format originated in the early 1980s.

Cited: *Akanthos Capital Management, LLC v. CompuCredit Holdings Corp.*, 2014 WL 896743 (N.D. Ga. Mar. 7, 2014).

Cited: *Country Vintner of North Carolina, LLC v. E. & J. Gallo Winery, Inc.*, 718 F.3d 249 (4th Cir. Apr. 29, 2013).

Cited: *E.E.O.C. v. SVT, LLC*, 2014 U.S. Dist. LEXIS 50114 (N.D. Ind. Apr. 10, 2014).

Cited: *Race Tires America, Inc. v. Hoosier Racing Tire Corp.*, 2012 WL 887593 (3d Cir. Mar. 16, 2012).

Cited: *Saliga v. Chemtura Corp.*, 2013 WL 6182227 (D. Conn. Nov. 25, 2013).

Cited: *In re Seroquel Products Liability Litigation*, 244 F.R.D. 650 (M.D. Fla., Aug. 21, 2007).

Cited: *Williams v. Sprint/United Management Company*, 230 F.R.D. 640 (D. Kan. 2005).

TIFF Group III: A one-dimensional compression format for storing black and white images that is utilized by many fax machines. See TIFF.

TIFF Group IV: A two-dimensional compression format for storing black and white images. Typically compresses at a 20-to-1 ratio for standard business documents. See TIFF.

Time Zone Normalization: See Normalization.

Toggle: A switch (which may be physical or virtualized on a screen) that is either on or off, and reverses to the opposite when selected.

Tone Arm: A device in a computer that reads to/from a hard drive.

Tool Kit Without An Interesting Name (TWAIN): A universal toolkit with standard hardware/software drivers for multi-media peripheral devices. Often used as a protocol between a computer and scanners or image capture equipment.

Toolbar: The row of graphical or text buttons that perform special functions quickly and easily.

Topology: The geometric arrangement of a computer system. Common topologies include a bus (network topology in which nodes are connected to a single cable with terminators at each end), star (LAN) designed in the shape of a star, where all end points are connected to one central switching device, or hub), and ring (network topology in which nodes are connected in a closed loop; no terminators are required because there are no unconnected ends). Star networks are easier to manage than ring topology.

Track: Each of the series of concentric rings contained on a hard drive platter.

Transmission Control Protocol/Internet Protocol (TCP/IP): The first two networking protocols defined; enable the transfer of data upon which the basic workings of the features of the Internet operate. See Internet Protocol, Port.

Trojan: A malware program that contains another hidden program embedded inside it for the purpose of discretely delivering the second program to a computer or network without the knowledge of the user or administrator. See Malware.

True Resolution: The true optical resolution of a scanner is the number of pixels per inch (without any software enhancements).

Twiki: A WikiWiki – Enables simple form-based Web applications without programming, and granular access control (though it can also operate in the classic ‘no authentication’ mode). Other enhancements include configuration variables, embedded searches, server-side includes, file attachments, and a plug-in API that has spawned over 150 plug-ins to link into databases, create charts, sort tables, write spreadsheets, make drawings, and so on.

Typeface: A specific size and style of type within a family. There are many thousands of typefaces available for computers, ranging from modern to decorative.

User Datagram Protocol (UDP): A protocol allowing computers to send short messages to one another. See Port.

Ultrafiche: Microfiche that can hold 1,000 documents/sheet as opposed to the normal 270.

Unallocated Space: The area of computer media, such as a hard drive, that does not contain readily accessible data. Unallocated space is usually the result of a file being deleted. When a file is deleted, it is not actually erased but is simply no longer accessible through normal means. The space that it occupied becomes unallocated space, i.e., space on the drive that can be reused to store new information. Until portions of the unallocated space are used for new data storage, in most instances, the old data remains and can be retrieved using forensic techniques.

Under-Inclusive: When referring to data sets returned by some method of query, search, filter or cull, results that are returned incomplete or too narrow. See False Negative.

Unicode: A 16-bit ISO 10646 character set accommodating many more characters than the ASCII character set. Created as a standard for the uniform representation of character sets from all languages. Unicode supports characters 2 bytes wide. Sometimes referred to as “double byte language.” See www.unicode.org for more information.

Uniform Resource Locators (URL): The addressing system used in the World Wide Web and other Internet resources. The URL contains information about the method of access, the server to be accessed and the path of any file to be accessed. Although there are many different formats, a URL might look like this:
http://thesedonaconference.org/publications_html. See Address.

Unitization – Physical and Logical: The assembly of individually scanned pages into documents. Physical Unitization utilizes actual objects such as staples, paper clips and folders to determine pages that belong together as documents for archival and retrieval

purposes. Logical unitization is the process of human review of each individual page in an image collection using logical cues to determine pages that belong together as documents. Such cues can be consecutive page numbering, report titles, similar headers and footers and other logical indicators. This process should also capture document relationships, such as parent and child attachments. See Attachment, Document or Document Family, Load File, Message Unit.

Cited: *Race Tires America, Inc. v. Hoosier Racing Tire Corp.*, 2012 WL 887593 (3d Cir. Mar. 16, 2012).

Universal Serial Bus (USB) Port: A port on a computer or peripheral device into which a USB cable or device can be inserted – quickly replacing the use or need for serial and parallel ports as it provides a single, standardized way to easily connect many different devices. See Flash Drive and Port.

UNIX: A software operating system designed to be used by many people at the same time (multi-user) capable of performing multiple tasks or operations at the same time (multi-tasking); common operating system for Internet servers.

Unstructured Data: Refers to free-form data which either does not have a data structure or has a data structure not easily readable by a computer without the use of a specific program designed to interpret the data; created without limitations on formatting or content by the program with which it is being created. Examples include word processing documents or slide presentations.

Upgrade: A newer version of hardware, software or application

Upload: To move data from one location to another in any manner, such as via modem, network, serial cable, internet connection or wireless signals; indicates that data is being transmitted to a location from a location. See Download.

User Created Metadata: Information about a file that is created and applied to a file by a user. Information includes the addressees of an email, annotations to a document and objective coding information. See Metadata.

Cited: *CBT Flint Partners, LLC v. Return Path, Inc.*, 737 F.3d 1320 (Fed. Cir. Dec. 13, 2013).

UTC: See Coordinated Universal Time.

UTF-8: A character encoding form of Unicode that represents Unicode code points with sequences of one, two, three or four bytes. UTF-8 can encode any Unicode character. It is the most common Unicode encoding on the Web and the default encoding of XML. An important advantage of UTF-8 is that it is backward compatible with the ASCII encoding, which includes the basic Latin characters. Consequently, all electronic text in the ASCII encoding is conveniently also Unicode. This backward compatibility was a primary reason for the invention of UTF-8. See ASCII, Unicode, UTF-16.

UTF-16: A character encoding form of Unicode that represents Unicode code points with sequences of one or two 16-bit code units. UTF-16 can encode any Unicode character. It is much less often used for data interchange than the UTF-8 encoding form. UTF-16 is commonly used in computer programming languages and programming APIs and is the

encoding used internally for file names by Microsoft Windows and NTFS. See Unicode, UTF-8.

Validate: In the context of this document, to confirm or ensure well-grounded logic, and true and accurate determinations.

Vector: Representation of graphic images by mathematical formulas. For instance, a circle is defined by a specific position and radius. Vector images are typically smoother than raster images.

Verbatim Coding: Manually extracting information from documents in a way that matches exactly as the information appears in the documents. See Coding.

Version, Record Version: A particular form or variation of an earlier or original record. For electronic records the variations may include changes to file format, metadata or content.

Vertical De-Duplication: A process through which duplicate Electronically Stored Information, as determined by matching hash values, are eliminated within a single custodian's data set. See Content Comparison, File level Binary Comparison Horizontal De-Duplication, Metadata Comparison, Near Duplicates.

Video Display Terminal (VDT): Generic name for all display terminals.

Video Electronics Standards Association (VESA): Concentrates on computer video standards.

Video Graphics Adapter (VGA): A computer industry standard, first introduced by IBM in 1987, for color video displays. The minimum dot (pixel) display is 640 by 480 by 16 colors. Then "Super VGA" was introduced at 800 x 600 x 16, then 256 colors. VGA can extend to 1024 by 768 by 256 colors. Replaces EGA, an earlier standard and the even older CGA. Newer standard displays can range up to 1600 by 1280.

Video Scanner Interface: A type of device used to connect scanners with computers. Scanners with this interface require a scanner control board designed by Kofax, Xionics or Dunord.

Virtual Private Network (VPN): A secure network that is constructed by using public wires to secure connect nodes. For example, there are a number of systems that enable creation of networks using the Internet as the medium for transporting data. These systems use encryption and other security mechanisms to ensure that only authorized users can access the network and that the data cannot be intercepted.

Virtualization: Partitioning a server into multiple virtual servers, each capable of running an independent operating system and associated software applications as though it were a separate computer. Virtualization is particularly useful for centralized IT infrastructures to manage multiple computing environments with the same set of hardware, and for cloud computing providers to provide customized interfaces to clients without investing in separate machines, each with its own operating system.

Virus: A self-replicating program that spreads on a computer or network by inserting copies of itself into other executable code or documents. A program into which a virus has inserted itself is said to be infected, and the infected file (or executable code that is not part

of a file) is a host. Viruses are a kind of malware that range from harmless to destructive and damage computers by either destroying data or overwhelming the computer's resources. See Malware.

Vital Record: A record that is essential to the organization's operation or to the re-establishment of the organization after a disaster.

Visualization: The process of graphically representing data.

Voice over Internet Protocol (VoIP): Telephonic capability across an Internet connection.

Volume: A specific amount of storage space on computer storage media such as hard drives, floppy disks, CD-ROM disks, etc. In some instances, computer media may contain more than one volume, while in others a single volume may be contained on more than one disk.

Volume Boot Sector/Record: When a partition is formatted to create a volume of data, a volume boot sector is created to store information about the volume. One volume contains the operating system and its volume boot sector contains code used to load the operating system when the computer is booted up. See Partition.

WAV: File extension name for Windows sound files.

Webmail: Email service that is provided through a website. See Email.

Website: A collection of Uniform Resource Indicators (URIs), including Uniform Resource Locators (URLs), in the control of one administrative entity. May include different types of URIs (e.g., FTP, telnet or Internet sites). See URI and URL.

What You See Is What You Get (WYSIWYG): Display and software technology that shows on the computer screen exactly what will print.

Wide Area Network (WAN): Refers generally to a network of PCs or other devices, remote to each other, connected by electronic means, such as transmission lines. See Network.

WiFi (Wireless Fidelity): Wireless networking technology that allows electronic devices to connect to one another and the Internet from a shared network access point.

Wiki: A collaborative website that allows visitors to add, remove, and edit content.

Wildcard Operator: A character used in text-based searching that assumes the value of any alphanumeric character, characters, or in some cases, words. Used to expand search terms and enable the retrieval of a wider range of hits.

Windows-1252: Also called ANSI, Western European and CP1252 (Microsoft code page 1252). A character encoding of the Latin alphabet used for most Western European languages. Windows-1252 is a superset of the ASCII and ISO 8859-1 standard character encodings. The characters that are included in Windows-1252, but that are not included in ISO 8859-1, are often the source of character interpretation and display problems in text on the Web and in electronic mail. Similar problems sometimes occur when text in the Windows-1252 encoding is converted to the UTF-8 encoding form of Unicode because UTF-8 is not wholly backward compatible with Windows-1252. The name ANSI is a

misnomer resulting from historical happenstance, but it is not incorrect to use it in contexts where its meaning is readily understood. See ASCII, ISO 8859-1.

Wireless Router: A hardware device that opens access to an Internet connection or network to a secured or unsecured connection via a receiver on a computer or other piece of hardware such as a printer permitting wireless transmission. See WiFi.

Workflow: The automation of a business process, in whole or part, during which Electronically Stored Information or tasks are passed from one participant to another for action according to a set of procedural rules.

Workflow, Ad Hoc: A simple manual process by which documents can be moved around a multi-user review system on an as-needed basis.

Workflow, Rule-Based: A programmed series of automated steps that route documents to various users on a multi-user review system.

Workgroup: A group of computer users connected to share individual talents and resources as well as computer hardware and software – often to accomplish a team goal.

Worm: A self-replicating computer program, sending copies of itself, possibly without any user intervention. See Malware.

World Wide Web (WWW): A massive collection of hypertext documents accessed via the Internet using a browser. The documents, also known as Web Pages, can contain formatted text, audio and video files, and multimedia programs.

Write Once Read Many Disks (WORM Disks): A popular archival storage media during the 1980s. Acknowledged as the first optical disks, they are primarily used to store archives of data that cannot be altered. WORM disks are created by standalone PCs and cannot be used on the network, unlike CD-ROM disks.

X.25: A standard protocol for data communications, which has largely been replaced by less complex protocols including the Internet Protocol (IP).

XML, XRML: See Extensible Markup Language.

Yottabyte: 1,024 zettabytes. See Byte.

Zettabyte: 1,024 exabytes. See Byte.

ZIP: A common file compression format that allows quick and easy storage for transmission or archiving one or several files.

Zip Drive: A removable disk storage device developed by Iomega with disk capacities of 100, 250 and 750 megabytes.

Zone OCR: An add-on feature of imaging software that populates data fields by reading certain regions or zones of a document and then placing the recognized text into the specified field.



**MOVING THE LAW FORWARD
IN A REASONED & JUST WAY**

Copyright 2014, The Sedona Conference
All Rights Reserved.
Visit www.thesedonaconference.org