

# THE SEDONA CONFERENCE JOURNAL®

*V o l u m e 8 ❖ F a l l 2 0 0 7*

## ARTICLES

- The Future of U.S. Federal Antitrust Enforcement:  
Learning from Past & Current Influences *James Langenfeld &  
Daniel R. Shulman*
- Ancillary Restraints Doctrine After *Dagher* *Gregory J. Werden*
- Get-Out-of-Jail-Free Cards:  
Amnesty Developments in the U.S. & Current Issues *Constance K. Robinson*
- If it Takes Two to Tango, Do They Conspire?  
*Twombly* & Standards of Pleading Conspiracy *Richard A. Duncan  
& Brian S. McCormac*
- The Sedona Conference® Introduction to the Legal & Economic  
Issues at the Intersection of the Patent & Antitrust Laws *The Sedona Conference® WG4*
- Claim Construction & Implicit Definitions Based on the  
Specifications Since *Phillips* *Robert Fram, Laurie Adams  
& Tanya Mazur*
- Solving Hobson's Choice: Suggested Changes to Willfulness  
Law in the Wake of *Knorr-Bremse* & *Echostar* *Timothy J. Malloy  
& Merle S. Elliott*
- Updates on the Corporate Attorney-Client Privilege *David M. Brodsky*
- Protection of the Attorney-Client Privilege in Criminal Investigations *Barry M. Sabin  
& Matthew R. Lewis*
- The Status & Content of Solicitor-Client Privilege in Canada:  
Questions Still Unanswered *The Hon. Gilles Létourneau*
- Comparative Approaches to the Attorney-Client Privilege in  
the U.S., Canada, U.K. & EU *Lista M. Cannon*
- The Practical Implications of Proposed Rule 502 *Ashish Prasad & Vazantha Meyers*
- The Sedona Guidelines: Best Practices Addressing Protective  
Orders, Confidentiality & Public Access in Civil Cases *The Sedona Conference® WG2*

## ESI SYMPOSIUM

- The Sedona Conference® Commentary on the Use of Search  
& Information Retrieval Methods in E-Discovery *The Sedona Conference® WG1*
- Search & Information Retrieval Science *Herbert L. Roitblat*
- The Sedona Conference® Commentary on Email Management:  
Guidelines for the Selection of Retention Policy *The Sedona Conference® WG1*
- The TREC Legal Track: Origins & Reflections on the First Year *Jason R. Baron*
- Not Your Mother's Rule 26(f) Conference Anymore *Moze Cowper & John Rosenthal*

The  
Sedona  
Conference®

ANTITRUST LAW,  
COMPLEX LITIGATION



AND INTELLECTUAL  
PROPERTY RIGHTS

---

# THE SEDONA CONFERENCE JOURNAL®

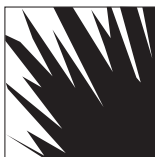
---

V O L U M E 8



F A L L 2 0 0 7

The  
sedona  
conference®



*The Sedona Conference Journal*<sup>®</sup> (ISSN 1530-4981) is published on an annual basis, containing selections from the preceding year's Conferences and Working Groups.

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 ([tsc@sedona.net](mailto:tsc@sedona.net)) or call (1-866-860-6600) to order or for further information. Complete Conference Notebooks for each of our Conferences are separately available. Check our website for further information about our Conferences, Working Groups, and publications:  
[www.thesedonaconference.org](http://www.thesedonaconference.org).

Comments (strongly encouraged) and requests to reproduce all or portions of this issue should be directed to:  
Executive Director, The Sedona Conference,  
180 Broken Arrow Way South, Sedona, AZ 86351-8998  
(toll free) 1-866-860-6600 or (tel.) 928-284-2698;  
fax 928-284-4240; email [tsc@sedona.net](mailto:tsc@sedona.net).

*The Sedona Conference Journal*<sup>®</sup> designed by  
Margo Braman of Studio D: [mbraman@sedona.net](mailto:mbraman@sedona.net).

Cite items in this volume to "8 Sedona Conf. J. \_\_\_\_ (2007)."

Copyright © 2007, The Sedona Conference.  
All Rights Reserved.

## PUBLISHER'S NOTE

---

Welcome to Volume Eight of *The Sedona Conference Journal*<sup>®</sup> (ISSN 1530-4981), an annual collection of articles originally presented at our conferences and papers prepared by our Working Groups over the past year. The Sedona Conference<sup>®</sup> 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 and several Working Groups 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 articles selected from our most recent Conferences on antitrust law and litigation (Fall 2006), patent litigation (Fall 2006), complex litigation (April 2007), and our Working Groups on Electronic Document Retention and Production (WG1), Protective Orders, Confidentiality and Public Access (WG2), and on the Intersection of the Patent and Antitrust Laws (WG4).

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 Groups, 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 to us electronically at [tsc@sedona.net](mailto:tsc@sedona.net), or by mail to The Sedona Conference, 180 Broken Arrow Way South, Sedona, AZ 86351-8998, USA. If you are interested in participating in one of our Regular Season Conferences (limited to 40 participants in addition to the 15-person faculty, to encourage the dialogue that is our hallmark), or in joining one of our Working Groups, please visit our website for further information ([www.thesedonaconference.org](http://www.thesedonaconference.org)).

*Richard G. Braman  
Executive Director  
The Sedona Conference  
Sedona, Arizona USA  
August 2007*

The Sedona Conference<sup>®</sup> gratefully acknowledges the substantial contributions of its Conference faculties, Working Group participants, members and observers, and our Advisory Board members, whose volunteer efforts make The Sedona Conference<sup>®</sup> a "thought-provoking and inspiring" experience providing content of immediate benefit to the Bench and Bar.

# THE SEDONA CONFERENCE® ADVISORY BOARD

---

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

**Tyler A. Baker, Esq.**  
Fenwick & West, Mountain View, CA

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

**Professor Stephen Calkins**  
Wayne State University Law School, Detroit, MI

**The Hon. Mr. Justice Colin Campbell**  
Federal Court of Appeal, Ottawa, Canada

**The Hon. John L. Carroll (Ret.)**  
Cumberland School of Law, Birmingham, AL

**Barbara Caulfield, Esq.**  
Affymetrix, Inc. Santa Clara, CA

**Joe Cecil, Ph.D., J.D.**  
Federal Judicial Center, Washington, DC

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

**The Hon. USMJ John M. Facciola**  
District of Columbia, Washington, DC

**Michael D. Hausfeld, Esq.**  
Cohen Milstein Hausfeld & Toll, Washington, D.C.

**Professor George A. Hay**  
Cornell Law School, Ithaca, NY

**The Hon. Katharine S. Hayden**  
District of New Jersey, Newark, NJ

**Ronald Hedges, Esq.**  
Nixon Peabody, New York, NY

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

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

**The Hon. Mr. Justice Gilles Letourneau**  
Federal Court of Appeal, Ottawa, Canada

**David H. Marion, Esq.**  
Montgomery McCracken Walker & Rhoads, Philadelphia, PA

**The Hon. J. Thomas Marten**  
District of Kansas, Wichita, KS

**Dianne M. Nast, Esq.**  
Roda & Nast, Lancaster, PA

**The Hon. USMJ Nan Nolan**  
Northern District of Illinois, Chicago, IL

**Raymond L. Ocampo, Jr., Esq.**  
Berkeley Center for Law & Technology, Berkeley, CA

**The Hon. Kathleen O'Malley**  
Northern District of Ohio, Cleveland, OH

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

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

**M. Laurence Popofsky, Esq.**  
Heller Ehrman, San Francisco, CA

**Jonathan M. Redgrave, Esq.**  
Redgrave Daley Ragan & Wagner, Washington, D.C.

**The Hon. James M. Rosenbaum**  
District of Minnesota, Minneapolis, MN

**Prof. Stephen A. Saltzburg, Esq.**  
GWU Law School, Washington, D.C.

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

**The Hon. USMJ Craig B. Shaffer**  
District of Colorado, Denver, CO

**Daniel R. Shulman, Esq.**  
Gray Plant Moory, Minneapolis, MN

**Robert G. Sterne, Esq.**  
Sterne Kessler Goldstein & Fox, Washington, D.C.

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

**Professor Jay Tidmarsh**  
Notre Dame Law School, Notre Dame, IN

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

**Craig W. Weinlein, Esq.**  
Carrington Coleman Sloman & Blumental, Dallas, TX

**The Hon. Carl J. West**  
LA Superior Court, Los Angeles, CA

# TABLE OF CONTENTS

---

<b>Publisher's Note</b> .....	<b>i</b>
<b>The Future of U.S. Federal Antitrust Enforcement: Learning from Past &amp; Current Influences</b> .....	<b>1</b>
<i>James Langenfeld &amp; Daniel R. Shulman</i>	
<b>Ancillary Restraints Doctrine After <i>Dagher</i></b> .....	<b>17</b>
<i>Gregory J. Werden</i>	
<b>Get-Out-of-Jail-Free Cards: Amnesty Developments in the U.S. &amp; Current Issues</b> .....	<b>29</b>
<i>Constance K. Robinson</i>	
<b>If It Takes Two to Tango, Do They Conspire? Twombly &amp; Standards of Pleading Conspiracy</b> .....	<b>39</b>
<i>Richard A. Duncan &amp; Brian S. McCormac</i>	
<b>The Sedona Conference® Introduction to the Legal &amp; Economic Issues at the Intersection of the Patent &amp; Antitrust Laws (Public Comment Version)</b> .....	<b>57</b>
<i>The Sedona Conference® WG4</i>	
<b>Claim Construction &amp; Implicit Definitions Based on the Specifications Since <i>Phillips</i></b> .....	<b>65</b>
<i>Robert Fram, Laurie Adams &amp; Tanya Mazur</i>	
<b>Solving Hobson's Choice: Suggested Changes to Willfulness Law in the Wake of <i>Knorr-Bremse</i> &amp; <i>EchoStar</i></b> .....	<b>77</b>
<i>Timothy J. Malloy &amp; Merle S. Elliott</i>	
<b>Updates on the Corporate Attorney-Client Privilege</b> .....	<b>89</b>
<i>David M. Brodsky</i>	
<b>Protection of the Attorney-Client Privilege in Criminal Investigations</b> .....	<b>105</b>
<i>Barry M. Sabin &amp; Matthew R. Lewis</i>	
<b>The Status &amp; Content of Solicitor-Client Privilege in Canada: Questions Still Unanswered</b> .....	<b>113</b>
<i>The Hon. Gilles Létourneau</i>	
<b>Comparative Approaches to the Attorney-Client Privilege in the U.S., Canada, U.K. &amp; EU</b> .....	<b>125</b>
<i>Lista M. Cannon</i>	
<b>The Practical Implications of Proposed Rule 502</b> .....	<b>133</b>
<i>Ashish Prasad &amp; Vazantha Meyers</i>	
<b>The Sedona Guidelines: Best Practices Addressing Protective Orders, Confidentiality &amp; Public Access in Civil Cases (Post-Public Comment Version)</b> ..	<b>141</b>
<i>The Sedona Conference® WG2</i>	
<b>ESI SYMPOSIUM</b>	
<b>The Sedona Conference® Best Practices Commentary on the Use of Search &amp; Information Retrieval Methods in E-Discovery (Public Comment Version)</b> .....	<b>189</b>
<i>The Sedona Conference® WG1</i>	
<b>Search &amp; Information Retrieval Science</b> .....	<b>225</b>
<i>Herbert L. Roitblat</i>	
<b>The Sedona Conference® Commentary on Email Management: Guidelines for the Selection of Retention Policy</b> .....	<b>239</b>
<i>The Sedona Conference® WG1</i>	
<b>The TREC Legal Track: Origins &amp; Reflections on the First Year</b> .....	<b>251</b>
<i>Jason R. Baron</i>	
<b>Not Your Mother's Rule 26(f) Conference Anymore</b> .....	<b>261</b>
<i>Moze Cowper &amp; John Rosenthal</i>	

# THE FUTURE OF US FEDERAL ANTITRUST ENFORCEMENT: LEARNING FROM PAST AND CURRENT INFLUENCES

---

*James Langenfeld<sup>1</sup> and Daniel R. Shulman<sup>2</sup>*

## I. INTRODUCTION<sup>3</sup>

There are a number of major influences that shape federal antitrust enforcement now, and in the future.<sup>4</sup> One of these influences is obviously who is running the Federal Trade Commission (FTC) and the Antitrust Division of the Department of Justice (DOJ), and their particular enforcement agendas. Much has been made of the continuity and general consensus about the role of antitrust, and in many ways this is true.<sup>5</sup> However, if there is a new administration or simply a change in the heads of the agencies, then past experience indicates that at least some change in emphasis will take place. Another direct influence on antitrust enforcement is what businesses decide to do. For example, if the number of mergers decline substantially, then presumably fewer mergers are likely to be challenged.

There are other influences that may be more indirect, but can have a longer run impact. The progress of economic thought has had a major influence on the agencies and courts over time, and that presumably will continue.<sup>6</sup> In addition, antitrust policy groups and associations can influence enforcement. For example, the Antitrust Modernization Commission (AMC) has recently made a number of recommendations.<sup>7</sup> These recommendations may guide the actions of the agencies, and could even lead Congress to enact changes to the law. Finally, the antitrust agencies success in winning cases is ultimately governed by the federal courts. Supreme Court and appeals court rulings set the bounds within the agencies and private plaintiffs can operate, as well as any legislative changes. Accordingly, the future of federal antitrust enforcement requires predicting the future of a variety of influences and how they will interact.

To project likely future enforcement, we perform five analyses. First, economists often look to past and current trends to predict the future; so we try to determine what those trends have been and what factors have influenced them. Second, we analyze some of the external market influences that could affect the level of antitrust enforcement. Third, we look at statements of those in charge of the agencies, taking into account that a new administration will occur in two years and will likely chose different senior officials at the agencies. Fourth, we look at the general state of economic thinking and research to see how this might influence future enforcement. Finally, we look at some key court decisions that are likely to limit and shape the success of federal enforcement efforts.

---

1 Director, LECG, and Adjunct Professor at Loyola University Law School, Chicago; former Director for Antitrust in the Federal Trade Commission's Bureau of Economics.

2 Principal, Gray, Plant, Moory, Moory & Bennett, Minneapolis, MN.

3 An earlier version of this article was presented at the Sedona Conference Antitrust Law and Litigation of October 26-27, 2006. The opinions expressed here are only those of the authors. However, we would like to thank Thomas Rosch, William Kovacic, and the participants in the Sedona Conference Antitrust Law and Litigation for their very helpful comments on the earlier version of this article, recognizing that this article does not necessarily reflect their opinions or the opinions of their colleagues. We also would like to thank Jamie Ziesch for her excellent research help.

4 The funding levels of the agencies can limit their level of activity, although with the advent of filing fees for mergers meeting the minimum size threshold there does not appear to be substantial budget pressure on either agency now.

5 See, for example, William E. Kovacic, "The Modern Evolution of U.S. Competition Policy Enforcement Norms", 71 *Antitrust L. J.* 377 (2004) and Thomas B. Leary, "The Essential Stability of Merger Policy in the United States", 70 *Antitrust L. J.* 105 (2002) (Hereafter Leary (2002)).

6 See, for example, William Kovacic, "The Intellectual DNA of Modern U.S. Competition Law for Dominant Firm Conduct: The Chicago/Harvard Double Helix," 1 *Columbia Law Review* 1-80 (2007). (Hereafter Kovacic (2007a)).

7 Antitrust Monopolization Commission, Report and Recommendations, April 2, 2007, available at [www.amc.gov/index.html](http://www.amc.gov/index.html). (Hereafter AMC (2007)).

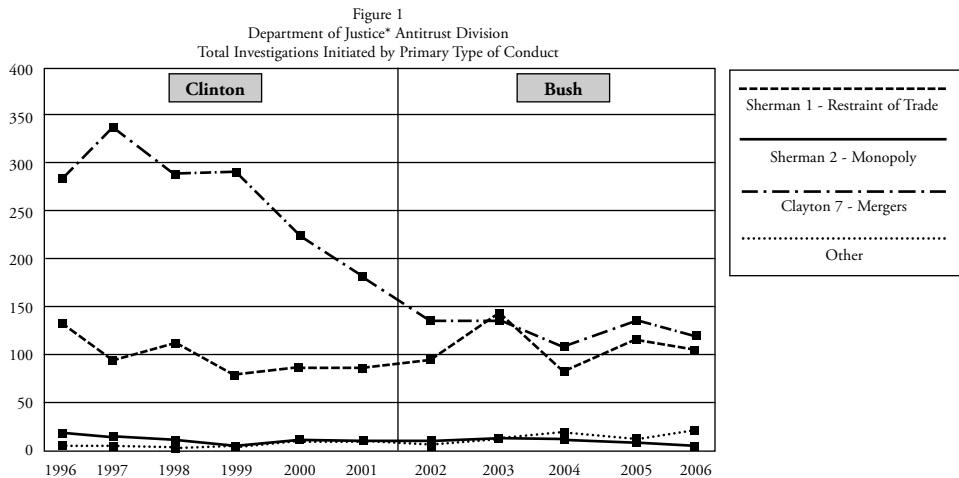
We do not analyze other potentially important influences in this article, but mention some of them here. In particular, we do not try to predict the influence of the various interested antitrust groups on future federal antitrust enforcement, except to discuss some of the AMC's recommendations. Other potentially important groups include the Antitrust Section of the American Bar Association (ABA), the International Bar Association, the National Association of Attorneys General, and the American Antitrust Institute. The ABA tends to argue caution in enforcing antitrust laws, while the latter two organizations have tended to argue for more active enforcement.

The remainder of the article begins with a discussion of the future of merger enforcement. We then discuss price fixing, cartels, and agreements among competitors; unilateral behavior and vertical non merger cases (monopolization, tying, bundling, predatory behavior, resale price maintenance, etc.); and finally a few other areas where there has been relatively little recent enforcement (Robinson-Patman, invitations to collude, facilitating practices, etc.).

## II. MERGER ENFORCEMENT

The agencies have traditionally devoted about half of their antitrust resources to merger enforcement since the implementation of the Hart-Scott-Rodino Act in 1976 premerger review process. Both agencies have been active in investigating and challenging mergers, and the vast majority of these have been mergers between direct competitors. Looking at the recent trend in the aggregate number of merger challenges and investigations can provide some insights into the level of enforcement activity that is likely to take place in the future.<sup>8</sup>

The DOJ publishes fairly detailed statistics on the number of its investigations and challenges by type of case, and the FTC makes less detailed information available in some of its reports.<sup>9</sup> Figure 1 shows the number of DOJ investigations of Sherman 1 restraint of trade, Sherman 2 monopolization, and Clayton 7 mergers from 1996 to 2006. Focusing on the merger investigations, it is clear that the number of investigations have fallen from its high during this period of 338 in 1997 to between less than 140 per year from 2002 to 2006.

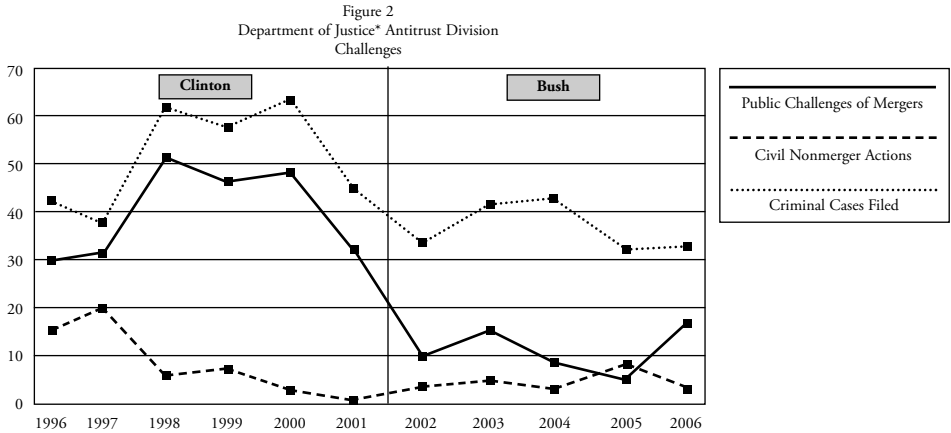


8 Clearly all mergers are not equal in terms of size, complexity, or the potential impact on consumers, so simple counts of the number of enforcement actions and investigations can be misleading. Moreover, as then FTC General Counsel, now Commissioner, William Kovacic has written "Beyond cases, the successful competition agency invests in research, holds hearings and workshops, performs empirical work, publishes studies, and submits advocacy comments to other public authorities." William E. Kovacic, "The Future of U.S. Competition Policy," September 2004, The Antitrust Source, [www.antitrustsource.com](http://www.antitrustsource.com). However, when the agencies report their enforcement activity they usually begin with the aggregate number of enforcement actions taken, suggesting the agencies put some weight on "the numbers". In addition, other researchers have analyzed the number of actions brought by the agencies to evaluate the level of federal antitrust enforcement activity. For example, Jonathan Baker and Carl Shapiro have recently studied merger challenges as a percentage of HSR filings since 1982 in their analysis of trends in merger enforcement at the FTC and DOJ. See Baker and Shapiro, "Reinvestigating Horizontal Merger Enforcement", April 10, 2007, available at <http://faculty.haas.berkeley.edu/shapiro/mergerpolicy.pdf>. (Referred to as Baker and Shapiro (2007) hereafter.) We caution putting too much weight on the number of challenges and investigations in any given year, since the mix of mergers presented to the agencies in any year may be substantially different from the norm.

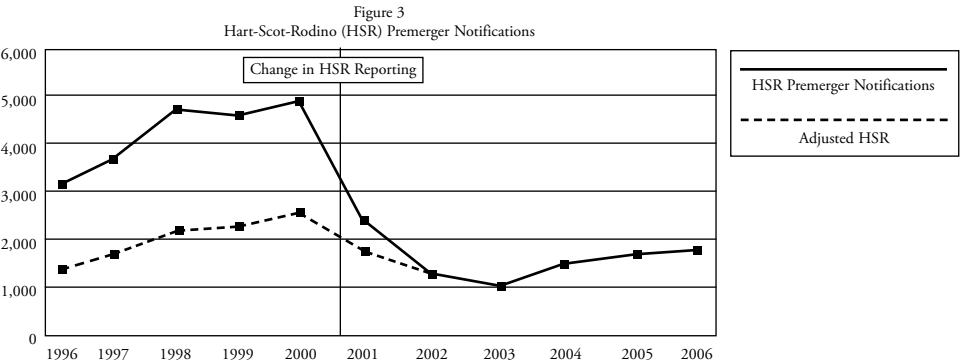
9 "Antitrust Division Workload Statistics FY 1996-2005," [www.usdoj.gov/atr/public/workstats.htm](http://www.usdoj.gov/atr/public/workstats.htm).



Figure 2 shows the number of DOJ public challenges to mergers for the same time period, as well as the number of criminal and civil nonmerger cases. During this time period, merger challenges by the DOJ were at their highest level from 1998 to 2000, ranging from 51 to 46 per year. The number of challenges fell sharply in the period 2002 to 2006, and ranging from a low of 4 challenges in 2005 to a high of 16 in 2006.

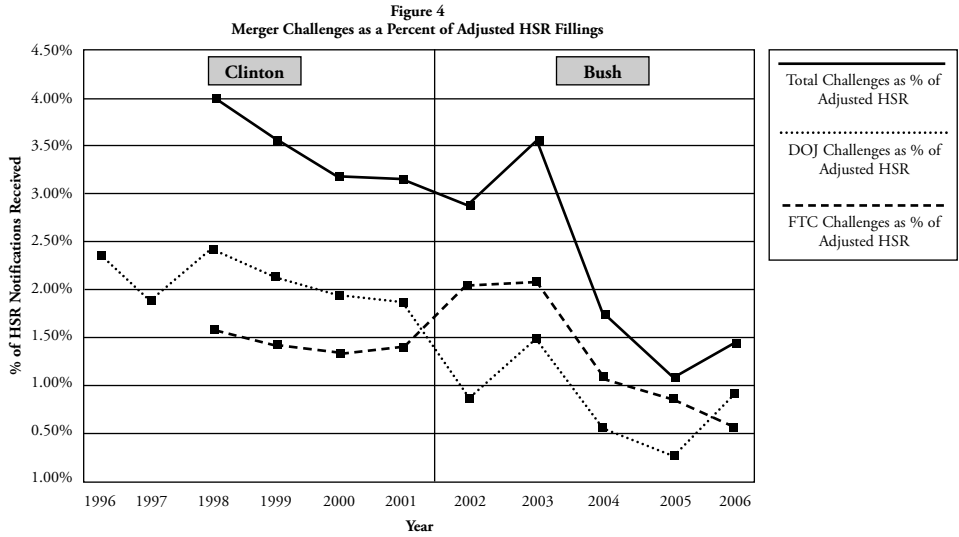


One reason for the fall in the number of investigations appears to be the drop off in the number of mergers after 2000, although measuring the fall off in the numbers is complicated by February of 2001 change in the Hart-Scott-Rodino (HSR) reporting requirements that raised the minimum cut off for reporting from \$15 million to \$50 million.<sup>10</sup> This change explains to a large degree the number of filings dropping from 4,926 to 2,376 from 2000 to 2001, as shown by the solid line in Figure 3. There is limited publicly available information that adjusts for this change over time in a way to use the HSR reporting as a measure of aggregate merger activity. However, in 2005 the FTC printed a chart in one of its annual reports that shows (without reporting the exact data) the number of transactions that would have been filed subject to the thresholds established in 2001.<sup>11</sup> We have roughly estimated these numbers from 1996 to 2001 based on that chart, and they are represented by the dashed line in Figure 3. These adjusted data also show a substantial (although less dramatic) drop in the number of mergers under the new HSR reporting requirements from a high of about 2,500 in 2000 to a low of about 1000 in 2003.



10 Premerger Notification, 66 Fed. Reg. 8680 (Feb. 1, 2001).  
 11 Federal Trade Commission, "The FTC in 2005: Standing up for Consumers and Competition," April 2005, p. 5, available at <http://www.ftc.gov/os/2005/04/0504abareportfinal.pdf>.

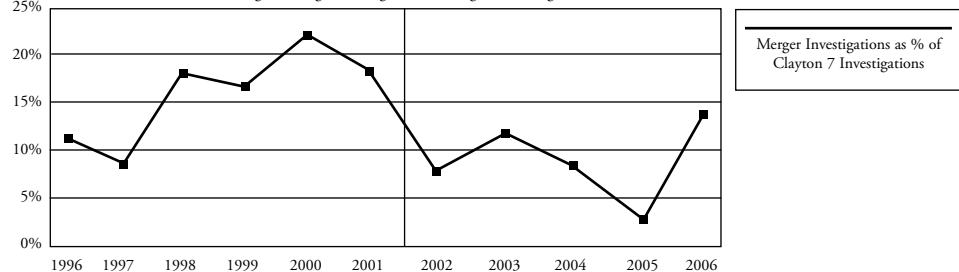
To control for this drop in merger activity, we calculate the percentage of the adjusted HSR filings received that were subject to DOJ public challenges, as illustrated in Figure 4. Although there are variations from year to year, the challenges as a percent of the adjusted HSR filings averaged about 2.2% per year from 1996 to 2001. During the years of 2002 to 2006, this average fell to about 1%.



The unadjusted data since 2002 is indicative of changes in the level of merger activity over that period. As shown in Figure 3, the number of filings after the reporting change fell to a low of 1,014 in 2003, and then rose to 1,768 in 2006, reflecting a reduction and then rise in merger activity over the last 6 years.<sup>12</sup> DOJ challenges as the percentage of HSR filings fell as the number of mergers increases. As illustration in Figure 4, in 2003 this percentage was 1.5%, when there were the fewest HSR filings. In 2004 this percentage fell to 0.62% and to 0.24% in 2005, before increasing to 0.9% in 2006.

Based on the adjusted and unadjusted data, the reduction in merger activity in 2001 through 2003 does not appear to explain all of the reduction in DOJ merger challenges since 2002. DOJ public challenges of mergers as a percentage of DOJ investigations should not be affected by the change in HSR reporting, and could reflect changes in DOJ management decision criteria for challenge mergers. Figure 5 shows the percentage of DOJ merger investigations that resulted in merger challenges. This percentage fell from an average of 19 percent per year for 1998 to 2000 to about 8 percent per year in the 2002 to 2005 period, before increasing to 14% in 2006. It appears that not only are there substantially fewer DOJ investigations in recent years, but the rate of challenges coming out of those investigations fell compared to the earlier period by more than half until 2006.

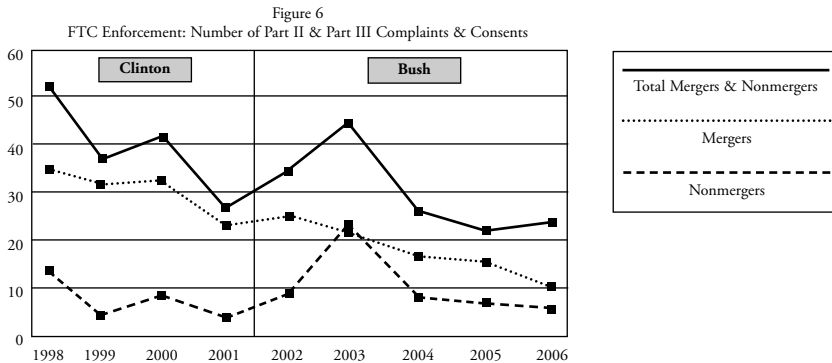
**Figure 5**  
Department of Justice\* Antitrust Division  
Percentage of Merger Investigations Leading to Challenges



12. Obviously using HRS filings for 2001 number will overstate the level of merger activity in that year, since two months of filings reflect the lower threshold prior to the change.

In general, these data suggest that merger enforcement activity at the DOJ has been reduced – rightly or wrongly. The upswing in merger challenges during 2006 has not raised the level of challenges or investigations to even half of their lowest levels from 1996 to 2001, and the fall in the number of mergers in 2002 and 2003 does not explain all of this drop. Presumably something substantial would need to change to reverse the trend in the foreseeable future.

The FTC does not publish similarly detailed systematic data on the number of investigations as far as our research has been able to determine, but one can track the number of merger related FTC preliminary injunctions, consents, Part III administrative complaints, and mergers abandoned due to FTC actions.<sup>13</sup> Figure 6 shows the sum of these actions from 1998 to 2005. The FTC's merger challenges have steadily fallen during this time period, from 34 in 1998 to 10 in 2006. Until 2006, the fall in the number of FTC enforcement actions had not been as substantial as the DOJ's, but there is a downward trend in merger challenges.



As illustrated in Figure 4 above, the FTC merger challenges as a percentage of the total adjusted HSR filings per year averaged about 1.4 percent both before and after 2002. This percentage was about 2 percent in the two years where there were the fewest filings, 2002 and 2003. It fell in 2004 to 2006 to less than half of the level in 2002 and 2003, also suggesting fewer FTC challenges in the last three years after taking into consideration the increase in aggregate merger activity during those three years.

Figure 4 also shows the combined percentage of DOJ and FTC challenges as a percent of adjusted HSR filings. There does not appear to be any significant drop in the total percentage of challenges through 2003. However, there has been a noticeable drop in this percentage in 2004 to 2006 as the number of mergers has increased.

In attempting to predict the level of future merger enforcement based on what we have observed in the recent past, the relevant question is why does there appear to have been so many investigations and challenges in late 1990s and relatively few recently? Examining Figures 1 to 6, one thing is apparent from the data. There are fewer investigations and challenges of mergers by both agencies since President Bush took office and installed new senior administrators at both agencies, and the reductions at least since 2003 do not appear to be due to fewer mergers. To illustrate the change in the administrations, we included vertical lines in Figures 1, 2, 5, and 6 to show the transition from the Clinton to the Bush administration enforcement between 2001 and 2002. This demarcation reflects what has become the standard analysis of treating the first year after a new administration as being more reflective of the prior than the new administration.<sup>14</sup>

<sup>13</sup> "Summary of Bureau of Competition Activity Fiscal Year 1998 through March 31, 2002," ABA Antitrust Spring Meeting, <http://www.ftc.gov/reports/aba/abaspring2002.pdf>; "Summary of Bureau of Competition Activity Fiscal Year 2002 through March 15, 2006," ABA Antitrust Spring Meeting, [www.ftc.gov/reports/aba/abaspring2006.pdf](http://www.ftc.gov/reports/aba/abaspring2006.pdf). Note that the FTC's 2004 and 2005 data as published do not reflect mergers abandoned due to FTC actions, and I have adjusted these figures based on discussions with FTC staff and "Annual Report to Congress Fiscal Year 2005," FTC and DOJ, [www.ftc.gov/reports/hsr05/P989316tweentyeighthannualhsrreport.pdf](http://www.ftc.gov/reports/hsr05/P989316tweentyeighthannualhsrreport.pdf), p. 13; "Annual Report to Congress Fiscal Year 2004," FTC and DOJ, <http://www.ftc.gov/reports/hsr04/5050810hsrrpt.pdf>, p. 11; "Annual Report to Congress Fiscal Year 2003," FTC and DOJ, <http://www.ftc.gov/os/2004/09/040903hsrrpt03.pdf>, p. 13.

<sup>14</sup> See, for example, Leary (2002); James Langenfeld and Louis Silvia "Federal Trade Commission Horizontal Restraint Cases: An Update," *The Antitrust Bulletin* 49, no.3, (Winter 2004) (hereafter Langenfeld and Silvia (2004)), pp. 521-591; and Baker and Shapiro (2007).

Although (as we discuss below) there are other reasons that at least in part likely explain the substantial decline, the data are consistent with the Bush administration reducing the level of enforcement actions. It is possible that more or less mergers will be challenged in the next two years based on the specific agendas of the current senior officials<sup>15</sup> and other factors external to the agencies. Clearly the DOJ substantially increased the number of mergers it challenged from 4 in 2005 to 16 in 2006. However, to the extent the lower number of challenges and investigations reflect Bush administration policy, there will presumably not be an increase to the levels seen in the Clinton administration – again rightly or wrongly.

Forecasting the level of enforcement beyond two years will obviously depend on who becomes President and who they will choose to run the agencies. It may be tempting to assume that if a Democrat is elected, then there will be more aggressive merger enforcement. In fact, the non Republicans FTC commissioners have been more aggressive about pursuing some mergers. Commissioners Jon Liebowitz and Pamela Harbour Jones dissented over closing the FTC's investigation of the Time Warner/Comcast acquisition and "swap" of certain Adelphia cable systems because "this transaction may raise the cost of sports programming to rival content distributors, and thus substantially lessen competition and harm consumers."<sup>16</sup> Moreover, Commissioner Harbour would have continued to litigate the FTC's district court loss in *Arch Coal* in an FTC Part III administrative proceeding.<sup>17</sup> On the DOJ side, the Antitrust Division recently approved the merger of AT&T and BellSouth without any divestitures or conditions, but Democratic members of the Federal Communications Commission raised substantial questions about whether the merger will harm competition.<sup>18</sup> Nevertheless, Republican administrations in the past have supported active antitrust enforcement. If there is a Republican president after the next election, the impact on the level of merger investigations and challenges will presumably depend on whether they share the same views of enforcement as President Bush and his administrators.

As discussed above, there are other factors that can affect the level of enforcement and the types of mergers challenged, so it would be too simplistic to attribute the fall off in the number of merger challenges merely to the change in administrations.<sup>19</sup> For example, both the FTC and DOJ have recently lost high profile merger challenges. The DOJ's unsuccessful challenge of the *Oracle/Peoplesoft*<sup>20</sup> merger is one, and the FTC's defeat in *Arch Coal*<sup>21</sup> is another. The FTC just received another defeat in district court in its attempt to challenge a merger in the petroleum industry.<sup>22</sup> At least some courts appear to be changing the old presumptions of *Philadelphia National Bank*<sup>23</sup> that a substantial increase in market share is the key for the agencies to successfully challenge a merger. The Supreme Court appears content with the situation, since it has not taken a Section 7 case in decades. Beyond what the agencies' policy priorities may be, the agencies are pragmatic. It is not surprising that such losses would make them cautious about bringing new cases that could establish even higher burdens of proof for future challenges. Interestingly, FTC Commissioner Pamela Harbour Jones' dissent on the Commission's decision not to challenge the Arch Coal merger in Part III highlights the potential problems of challenging mergers and losing. She states "[i]f the Commission does not continue its enforcement action, we run the risk that the district court opinion will impose an unnecessarily high burden of proof for future merger challenges predicated on coordinated effects."<sup>24</sup> These court imposed limitations on merger enforcement suggest that an increase in the aggregate level of merger challenges will be difficult, even if the agencies want to expand their activity.

15 There is evidence that shows that administrators can have a substantial impact on policy. See, for example, Thomas F. Walton and James Langenfeld, "Regulatory Reform Under Reagan - The Right Way and the Wrong Way," in *Regulation and the Reagan Era: Politics, Bureaucracy and the Public Interest*, ed. Roger Meiners and Bruce Yandle (San Francisco: Independent Institute, 1989).

16 Statement of Commissioners Jon Liebowitz and Pamela Harbour Jones (Concurring in Part, Dissenting in Part), Time Warner/Comcast/Adelphia, Filed No. 051-0151 available at FTC web site.

17 Dissenting Statement of Commissioner Pamela Harbour Jones In The Matter of Arch Coal Inc., et. al., Docket No. 9316/File No. 031- 0191, available at FTC website.

18 See, for example, Paul Davidson, "BellSouth Deal Faces FCC Fight", USA TODAY, October 12, 2006, 3B.

19 Another possible explanation of the drop in the number of mergers challenged and investigated is the changing nature of the types of mergers over time. Clearly there have been many more private equity acquisitions in recent years. Such acquisitions are less likely to reduce competition than traditional strategic mergers, where there is a greater likelihood of competitors merging. To our knowledge, there are no reliable estimates of percentage of transaction that are private equity, so it is difficult to test the likely effect of these acquisitions on the number of potentially anticompetitive mergers. While there are estimates that total value of private equity investments has grown four fold from 2002 to 2005, the actual number of these transactions has fallen from 128 to 67 in this period. Apax Partners, Unlocking Global Value, Future Trends in Private Equity Investment World Wide, at 8. available at <http://www.apax.com>.

20 *United States v. Oracle Corp.*, 331 F.Supp.2d 1098, 1123 (N.D. Calif. 2004).

21 *Federal Trade Commission v. Arch Coal, Inc.*, 329 F.Supp. 2d 109 (D.D.C. 2004)

22 *Federal Trade Commission v. Paul L. Foster, Western Refining, Inc., and Giant Industries, Inc.* (United States District Court of New Mexico), No. CIV 070-352/JB/ACT (Slip. Op.), May 29, 2007.

23 *United States v. Philadelphia National Bank*, 374 U.S. 321 (1963).

24 Dissenting Statement of Commissioner Pamela Harbour Jones In The Matter of Arch Coal Inc., et. al., Docket No. 9316/File No. 031- 0191, at 1-2, available at FTC website.

The FTC has clearly tried to re-establish its role in merger enforcement in the face of adverse court decisions in at least one area, hospital mergers. The FTC and DOJ have lost all of the preliminary injunctions brought in hospital mergers since 1992, a total of six losses in a row.<sup>25</sup> Failing consistently in district courts since 1992, the FTC chose to challenge a consummated hospital merger of the Evanston-Northwestern and Highland Park hospitals in the northern Chicago suburbs with a Part III administrative complaint. The FTC sought to show that prices had increased as a result of the merger, and that the patient flow data that most of the recent hospital decisions have relied upon for geographic market definition was unreliable and misleading. The administrative law judge found that the merger was anticompetitive,<sup>26</sup> and the Commission is sitting in judicial review of the ALJ's decision at this writing. The final outcome of the case may determine whether the FTC and DOJ will be able to successfully challenge any hospital mergers in the future.

In recent years there have been very few merger challenges that do not involve direct competitors. Since 2000 the FTC has had at least two (publicly announced) investigations in which vertical issues were prominent – the proposed *Cytyc/Digene*<sup>27</sup> and *Avant!/Synopsis*<sup>28</sup> vertical mergers. The FTC challenged the first and the merger was abandoned, and it did not challenge the second. This is an area where the economics and case law is more complex, and there is much less consensus on when (and if) such mergers should be challenged. The agencies' policies on vertical mergers – other than they are much less likely to present competitive problems than horizontal mergers – are very unclear. This can be seen in the *Merger Guidelines*. In 1984, the *Merger Guidelines* covered vertical as well as horizontal mergers, but in 1992 the agencies only issued the Horizontal Merger Guidelines. A quick review of the 1984 *Guidelines* on vertical mergers show them sadly not up to date with current economic thinking or with the alternative theories of those vertical mergers that have been challenged or investigated. The AMC apparently realized this, since one of its recommendations was to update the *Merger Guidelines* to include how the agencies evaluate non-horizontal mergers.<sup>29</sup> The lack of any revision of the out-of-date 1984 *Guidelines* on vertical mergers suggest this has been an area of low priority, so the agencies seem quite unlikely to challenge many vertical mergers in the foreseeable future. Moreover, given the strong stance that the DOJ took on the European Commission's challenge of the GE/Honeywell merger based on a leveraging theory, it seems quite unlikely that many conglomerate mergers will be challenged absent a sea change in enforcement policy.<sup>30</sup>

The other longer run influences on merger enforce include advances in economic thinking and policy/legislative recommendations from antitrust experts and organizations. Clearly the "Chicago School" of economics has had a major influence on merger enforcement, which can be traced back to the revision of the *Merger Guidelines* in 1982. "Post Chicago School" thought, and in particular the introduction of game theory considerations, affected the 1992 revision of the *Horizontal Merger Guidelines* by shifting the focus of merger inquiries more to unilateral competitive effects from mergers, and to some degree de-emphasized concerns about coordinated effects. With the new administrators put in place by President Bush, the potential for coordinated effects from mergers regained more focus of the agencies' inquiries.

The net effect of recent economic thinking has led to a de-emphasis of reliance on simple market share calculations and an increased focus on an economic analysis of the likely competitive

25 See H.E. Frech III, James Langenfeld, and R. Forrest McCluer, "Elzinga-Hogarty Tests and Alternative Approaches for Market Share Calculations in Hospital Markets," *Antitrust Law Journal*, 71, no.3 (2004), pp. 921-947.

26 Available at <http://www.ftc.gov/os/adipros/49315/index.htm>

27 FTC Press Release, *Federal Trade Commission Seeks to Block Cytyc Corporation's acquisition of Digene Corporation*, File No. 021-0098 (June 24, 2002), available at [http://www.ftc.gov/opa/2002/06/cytyc\\_digene.htm](http://www.ftc.gov/opa/2002/06/cytyc_digene.htm).

28 See FTC Press Release, *Federal Trade Commission Votes to Close Investigation of Acquisition of Avant! Corporation by Synopsis, Inc.*, File No. 021-0049 (July 26, 2002), available at <http://www.ftc.gov/opa/2002/07/avant.htm>. As three Commissioners noted in their separate statements, the Commission intended to watch this market closely in the future, and did not rule out the possibility of seeking relief in the future if market effects prove to be more harmful than was apparent in advance of the merger. *Statement of Commissioner Thomas B. Leary, Synopsis Inc./Avant! Corp.*, File No. 021-0049 (July 26, 2002), available at <http://www.ftc.gov/os/2002/07/avantlearystmt.htm>; *Statement of Commissioner Mozelle W. Thompson, Synopsis Inc./Avant! Corp.*, File No. 021-0049 (July 26, 2002), available at <http://www.ftc.gov/os/2002/07/avantthompsonstmt.htm>; *Statement of Commissioner Sheila F. Anthony, Synopsis Inc./Avant! Corp.*, File No. 021-0049 (July 26, 2002), available at <http://www.ftc.gov/os/2002/07/avantanthonystmt.htm>. No subsequent public action has been taken by the FTC.

29 AMC 2007, p. 11.

30 See, for example, Deborah Platt Majoras, then Deputy Assistant Attorney General of the Antitrust Division, "GE-Honeywell: The U.S. Decision," Remarks before the Antitrust Section of the State Bar of Georgia, November 29, 2001, available at <http://www.usdoj.gov/atr/public/speeches/9893.htm>.

effects of the merger.<sup>31</sup> This can be seen in the agencies recent publications relating their merger enforcement. Clearly, the FTC's retrospective on factors influencing its challenges of mergers indicates that it only uses market concentration statistics (*i.e.*, HHIs) to decide which mergers to investigate, and will only challenge mergers well above the Horizontal Merger Guidelines' thresholds for "highly concentrated" markets.<sup>32</sup> Now there is more focus on the number of competitors (mergers of 3 to 2 or to monopoly typically challenged, relatively little chance of challenging 4 to 3 or higher) and potential entry. After the extensive use of econometric modeling of demand and game theory models to predict price increases from a merger in the late 1990s,<sup>33</sup> economic analysis now focuses more on "natural experiences" that try to predict the impact of a merger based on similar events in the industry or other geographic areas.<sup>34</sup> These shifts from a more structural to a more detailed economic approach to predicting the effects of mergers have occurred under both Democratic and Republican administrations. However, economists Jonathan Baker and Carl Shapiro have recently suggested that this trend may have gone too far, stating "that large increases in market concentration should be given real weight in merger analysis, and that any contrary presumption that 'two is enough' is unsupported."<sup>35</sup>

In general, the influence of economic thought on economic policy seems to occur with a lag, so looking at current economic thinking may be a good way to predict how the agencies will analyze mergers in the future. As illustrated by the court decisions mentioned above, it seems that courts may have not yet accepted some of the current economic thought on mergers, instead relying more on the ability of firms to reposition, requiring a great deal of evidence on precisely defined markets, and following a more traditional Chicago School skepticism of government intervention.<sup>36</sup>

### III. AGREEMENTS AMONG COMPETITORS

An area of substantial agreement by both legal scholars and economists is that agreements among competitors that restrict competition are the most dangerous area of potentially anticompetitive practices. Certain types of agreements, such as price fixing and bid rigging, have been investigated and challenged regularly by the DOJ. As can be seen in Figure 1, the DOJ has consistently investigated about 100 of these cases a year over the last decade. Although the number of investigations has remained relatively stable, the number of criminal cases filed has dropped from about 60 per year between 1998 and 2000 to an average of less than 40 per year from 2002 to 2006, as shown in Figure 2.<sup>37</sup>

These data suggest that the DOJ has consistently devoted substantial investigative resources to price fixing and bid rigging criminal cases. However, the data suggest that the DOJ has tightened its criteria for bringing these cases since 2002 compared to the last three full years of the Clinton administration.

Speeches by DOJ officials do not suggest any substantial change in criminal enforcement efforts,<sup>38</sup> nor does it seem likely that a change in administrations or new economic analyses will reduce Federal policy with regard to this core area of antitrust enforcement. There has been a trend in sharing information about international cartels between the competition agencies in different countries, which is likely to continue. Accordingly, it is more likely that international cartels will be detected and the DOJ should be more likely to challenge agreements involving larger markets. There is also some chance that economic analysis may play a larger role in cartel cases, but this has yet to be firmly established.<sup>39</sup>

31 See, for example, James A. Keyte and Neal R. Stoll, "Markets? We Don't Need No Stinking Markets!" *Antitrust Bulletin* 49, no. 3 (Fall 2004), p. 593; and Malcolm Coate, "Empirical Analysis of Merger Enforcement Under the 1992 Merger Guidelines," *Review of Industrial Organization*, 27, (Spring 2005), p. 279.

32 Staff Report of the Federal Trade Commission, "Horizontal Merger Investigation Data, 1996-2003," Revised August 31, 2004, available at FTC website.

33 Robert Lande and James Langenfeld, "From Surrogates to Stories: The Evolution of Federal Merger Policy," *Antitrust*, Spring 1997, 5-9.

34 See Mary Coleman and James Langenfeld, "Natural Experiments," in *Issues in Competition Law and Policy* (American Bar Association, forthcoming 2007).

35 Baker and Shapiro (2007), p. 22.

36 For discussion of other court decisions on mergers, see Baker and Shapiro (2007). In addition, these authors also performed an interesting survey of antitrust practitioners' opinions on merger enforcement.

37 The number of civil non-merger actions (which includes both Section 1 non-criminal competitor agreements and Section 2 monopolization cases) has been less than ten for every year since 1998.

38 See, for example, Gerald Masoudi, Deputy Assistant Attorney General Antitrust Division, "Cartel Enforcement in the United States (and Beyond)," Presented at the Cartel Conference, Budapest, Hungary, February 16, 2007, available at <http://www.usdoj.gov/atr/public/speeches/221868.htm>.

39 See *In re High Fructose Corn Syrup Antitrust Litigation*, 295 F.3d 651 (7th Cir. 2000) (Posner, J.) (use of economic evidence to prove collusion). For a brief discussion of the state of economic analysis' relevancy for inferring conspiracies, see James Langenfeld and James Morsch, "Refining the Matsushita Standard and The Role Economics Can Play", *Loyola University Chicago Law Journal*, Vol. 38, No. 3, Spring 2007, 507-512.

The FTC has generally taken the lead in challenging non-criminal agreements among competitors.<sup>40</sup> The vast majority of the FTC's non-merger enforcement efforts in the last 25 years have been challenging non-criminal agreements among competitors. For example, Figure 6 shows the number of FTC non-merger challenges by year from 1998 to 2006. Of the 23 non-merger enforcement actions in 2003, 19 were distinct horizontal restraints cases.<sup>41</sup> Almost half of the cases between 1993 and 2003 involved ambulatory health care or professional, scientific, or technical services. More recently, the FTC has pursued agreements between pharmaceutical firms, such as the so-call "reverse payments" cases where pioneer drug firms have paid money to generic drug firms to set or influence the generic entry date in the context of patent litigation settlements.<sup>42</sup>

With the exception of the 2003, which was an all time high in the FTC bringing agreements among competitors cases, there has been a fairly constant level of FTC enforcement actions in the past 10 years. Research shows that the more recent cases have tended to involve cases brought under traditional economic theories, such as attempts by sellers to fix prices, rather than "raising rivals costs" or "raising own cost" theories.<sup>43</sup>

To a large degree the FTC under the Clinton administration moved away from its approach to analyzing agreements among competitors under a "quick look" approach outlined in its 1980s *Mass. Board*<sup>44</sup> to more traditional per se treatment of agreements. In particular, the FTC took a per se approach in *California Dental Association*,<sup>45</sup> which was overturned by the appeals court for lack of sufficient market analysis of the impact of the challenged advertising restrictions. Under the Bush administration, the FTC has attempted to revitalize the *Mass. Board* approach in its decisions on the *Polygram*<sup>46</sup> ("Three Tenors") and *Schering Plough*<sup>47</sup> cases.<sup>48</sup>

Court decisions may be a major factor in limiting successful future challenges to agreements among competitors. For example, *Schering Plough* was a "reverse payments" case, where the FTC found there was a need for a full "rule of reason" case. However, the appeals court reversed the Commission decision of liability for lack of substantial evidence that the agreement delayed entry in the context of a patent settlement, which the FTC appealed to the Supreme Court. The Supreme Court denied *certiorari*, which raises questions about whether and how the FTC should be investigating and challenging such agreements. Moreover, the FTC appealed without the support of the DOJ, and the Supreme Court requested the opinion of the DOJ on the matter. The DOJ opposed granting *certiorari* in its submission, clearly showing a split between the agencies on these cases.

In addition to this set back to the FTC's challenges to non-criminal agreements not to compete, recent Supreme Court decisions are likely to make it more difficult for the agencies to challenge agreements not to compete. For example, *Texaco, Inc. v. Dagher*, 547 U.S. \_\_\_, 126 S.Ct. 1276, 164 L.Ed.2d 1 (2006) raises the question of whether joint ventures will become something of a safe harbor for concerted activity of competitors. If a combination must be challenged under the equivalent of rule of reason requirements of Section 7 of the Clayton Act because it claims to be a joint venture, then venture participants' conduct becomes difficult to challenge under Sections 1 or 2 of the Sherman Act, and perhaps under Section 5 of the FTC Act. *Dagher* appears to be a rather narrow decision, largely because the Court did not address the broader areas raised by the briefing. It appears to have left untouched the ancillary restraint doctrine of *Addyston Pipe*,<sup>49</sup> although one might legitimately question whether the doctrine has become inapplicable to a "fully integrated" joint venture. One can also question whether *Copperweld*<sup>50</sup> has been extended *sub silentio* to joint ventures.

40 James Langenfeld and Louis Silvia, "The Federal Trade Commission's Horizontal Restraint Cases: An Economic Perspective," *Antitrust Law Journal*, (Spring 1993), pp. 653-697; James Langenfeld and Louis Silvia "Federal Trade Commission Horizontal Restraint Cases: An Update," *The Antitrust Bulletin* 49, no.3, (Winter 2004) (hereafter Langenfeld and Silvia 2004), pp. 521-591. (Hereafter *Langenfeld and Silvia (2004)*.)

41 Langenfeld and Silvia (2004), p. 522.

42 See, for example, James Langenfeld and Wenqing Li, "Intellectual Property and Agreements to Settle Patent Disputes: The Case of Partial Settlement Agreement with Payments from Branded to Generic Drug Manufacturers," *Antitrust Law Journal* 70, Issue 3 (Spring 2003), pp. 777-818.

43 Langenfeld and Silvia (2004), p. 536.

44 Massachusetts Board of Registration in Optometry, 110 F.T.C. 549 (1988).

45 California Dental Association, 121 F.T.C. 190 (1996), *aff'd*, 128 F.3d 720 (9th Cir. 1997), *vacated, remanded* 526 U.S. 756 (1999), *rev'd, remanded* 224 F.3d 942 (9th Cir. 2000).

46 PolyGram Holding, Inc., Docket No. 9298, July 24, 2003, available at <http://www.ftc.gov/os/200307/polygramopinion.pdf>.

47 *FTC v. Schering-Plough Corp.*, 402 F.3d 1056 (11th Cir. 2005), *cert. denied*, 126 S.Ct. 2929 (2006).

48 Langenfeld and Silvia (2004).

49 *United States v. Addyston Pipe & Steel Co.*, 85 F. 271 (6th Cir. 1898), *modified and aff'd*, 175 U.S. 211 (1899).

50 *Copperweld Corp. v. Independence Tube Corp.*, 467 U.S. 752 (1984).

The *Copperweld* issue was extensively briefed, but not addressed in the opinion. The opinion also contained no discussion of *Citizen Publishing v. United States*,<sup>51</sup> the precedent on which the 9th Circuit and respondents principally relied.

In the future, *Dagher* will need to be distinguished from numerous joint venture precedents, including *BMI v. CBS*,<sup>52</sup> *Cal. Dental v. FTC*,<sup>53</sup> *FTC v. Indiana Fed. of Dentists*,<sup>54</sup> *NCAA v. Board of Regents*,<sup>55</sup> *Freeman v. San Diego Board of Realtors*,<sup>56</sup> and *Polygram Holding v. FTC*.<sup>57</sup> *Dagher* may also have implications for the application of the FTC's "quick look" analysis. The Third Circuit approved the FTC framework for analyzing joint venture conduct under the quick look in *Polygram*, and the question is whether *Dagher* will be read to affect that.<sup>58</sup>

In the recent Supreme Court term, there has been another potentially significant development in the law regarding concerted action, the Supreme Court's decision in *Bell Atlantic Corp. v. Twombly*.<sup>59</sup> This decision puts to rest any idea that vague and conclusory allegations of conspiracy premised on parallel conduct are sufficient to survive a motion to dismiss, eliminating the possibility that discovery will provide evidence in support of such allegations. The Court summarized its holding in *Twombly* as follows:

The question in this putative class action is whether a Section 1 complaint can survive a motion to dismiss when it alleges that major telecommunications providers engaged in certain parallel conduct unfavorable to competition, absent some factual context suggesting agreement, as distinct from identical, independent action. We hold that such a complaint should be dismissed.<sup>60</sup>

In *Twombly*, the plaintiff customer class alleged that the defendant Baby Bells or "incumbent local exchange carriers ('ILECs')" engaged in parallel refusals to deal with "competitive local exchange carriers ('CLECs')," with which the plaintiff class did business.<sup>61</sup> The plaintiffs further alleged that such parallel refusals to deal were evidence of a conspiracy among the defendants. To this, the Court responded,

[A] plaintiff's obligation to provide the 'grounds' of his 'entitle[ment] to relief' requires more than labels and conclusions, and a formulaic recitation of the elements of a cause of action will not do. . . . Factual allegations must be enough to raise a right to relief above the speculative level, *see* C. Wright & A. Miller, *Federal Practice and Procedure*, Section 1216, pp. 235-36 (3d ed. 2004) (hereinafter Wright & Miller) ("[T]he pleading must contain something more than . . . than . . . a statement of facts that merely creates a suspicion [of] a legally cognizable right of action") . . . .<sup>62</sup>

The Court accordingly held "that stating such a claim requires a complaint with enough factual matter (taken as true) to suggest that an agreement was made."<sup>63</sup> As to proceeding to

51 394 U.S. 131 (1969).

52 441 U.S. 1 (1979).

53 526 U.S. 756 (1999).

54 476 U.S. 447 (1986).

55 468 U.S. 85 (1984).

56 322 F.3d 1133 (9th Cir.), *cert. denied*, 540 U.S. 940 (2003).

57 416 F.3d 29 (D. C. Cir. 2005).

58 *See* James A. Keyte, "Dagher and 'Inside' Joint Venture Restraints," *Antitrust* 20, no. 3 (Summer 2006), p. 44.

59 — U.S. —, 127 S.Ct. 1955, 2007 WL 1461066 (May 21, 2007).

60 2007 WL 1461066 at \*4.

61 *Id.* at \*4-5.

62 *Id.* at \*8.

63 Thus, it is one thing to be cautious before dismissing an antitrust complaint in advance of discovery, *cf. Poller v. Columbia Broadcasting System, Inc.*, 368 U.S. 464, 473, 82 S.Ct. 486, 7 L.Ed.2d 458 (1962), but quite another to forget that proceeding to antitrust discovery can be expensive. As we indicated over 20 years ago in *Associated Gen. Contractors of Cal., Inc. v. Carpenters*, 459 U.S. 519, 528, n.17, 103 S.Ct. 897, 74 L.Ed.2d 723 (1983), "a district court must retain the power to insist upon some specificity in pleading before allowing a potentially massive factual controversy to proceed." *See also Car Carriers, Inc. v. Ford Motor Co.*, 745 F.2d 1101, 1106 (C.A.7 1984) ("[T]he costs of modern federal antitrust litigation and the increasing caseload of the federal courts counsel against sending the parties into discovery when there is no reasonable likelihood that the plaintiffs can construct a claim from the events related in the complaint") . . . . Probably, then, it is only by taking care to require allegations that reach the level suggesting conspiracy that we can hope to avoid the potentially enormous expense of discovery in cases with no "reasonably founded hope that the [discovery] process will reveal relevant evidence" to support a Section 1 claim. [*Id.*]



discovery, the Supreme Court was clear that the complaint must state a cognizable claim before discovery can be permitted.<sup>64</sup>

After *Twombly*, the question for future antitrust enforcement is whether actions seeking to prove conspiracy from circumstantial evidence will find their path much more difficult, particularly at the pleading stage. This may not affect the FTC's enforcement efforts, since the FTC can develop a detailed record if it chooses to pursue a case through an administrative complaint.

The future enforcement of non-criminal agreements among competitors cases appears mixed. For example, virtually all of these enforcement actions relating to the professions have resulted in settlement agreements, so the Commission's recent victory in the litigated *North Texas Specialty Physicians* matter is likely to add further weight to these cases – as long as that decision stands.<sup>65</sup> We can therefore expect the FTC to continue to challenge physician and other professional agreements cases. The FTC loss in *Schering Plough* and the DOJ's opposition to *certiorari* make the agency's future efforts to challenge agreements in pharmaceuticals where there are patent issues less clear. However, the current composition of the FTC, including the non Republican Commissioners, seems committed to future challenges to agreements among pharmaceutical companies.<sup>66</sup> More generally, the recent Supreme Court decisions are likely to discourage the agencies challenging from bringing some agreements among competitors cases.

#### IV. UNILATERAL ANTICOMPETITIVE BEHAVIOR AND VERTICAL RESTRAINTS

Unlike merger enforcement and challenges to agreements among competitors, the agencies' recent track record in challenging monopolistic practices and vertical restraints suggests that unless there is a substantial change in policy, court decisions, or empirical economic analysis, there will be very few enforcement actions in these areas. As can be seen in Figure 1, the number of DOJ Sherman Section 2 monopolization investigations has declined from 19 in 1996 to about half or less than that since 1998, with 3 such investigations in 2006. Unlike mergers, these investigations typically take so much time that any actions resulting from an investigation will not be reported until subsequent years. Not surprisingly then, the number of DOJ civil non-merger actions (which include Sherman Section 2 cases) has fallen from 20 in 1997 to between zero and 8 in any year since 1998, as shown in Figure 2. The most notable of the DOJ monopolization cases under the Clinton Administration was the *Microsoft*<sup>67</sup> case, which the DOJ won. However, there has been criticism by some that the consent obtained by the DOJ under the Bush Administration did not go far enough to remedy Microsoft's behavior, and the European Commission has taken a much more aggressive stance in limiting Microsoft's actions.<sup>68</sup> There have been no similar major monopolization cases filed by the DOJ since *Microsoft*. Moreover, as discussed below, a review of the DOJ *amicus* briefs to the Supreme Court in the last 10 years indicates the DOJ has consistently sided with the defendants' positions in these cases, rather than the plaintiffs.

The FTC also has not brought many monopolization or vertical restraints cases under either the Clinton or Bush Administrations. However, the FTC has devoted considerable time and effort in analyzing certain types of unilateral actions related to the enforcement intellectual property rights.

In *Rambus*<sup>69</sup> and *Unocal*<sup>70</sup>, the FTC separately found each of these companies had engaged in anticompetitive behavior in the context of forming industry standards. The FTC concluded in these

64 The Court continued, "Hence, when allegations of parallel conduct are set out in order to make a Section 1 claim, they must be placed in a context that raises a suggestion of a preceding agreement, not merely parallel conduct that could just as well be independent action. . . ."

An allegation of parallel conduct is thus much like a naked assertion of conspiracy in a Section 1 complaint: it gets the complaint close to stating a claim, but without some further factual enhancement it stops short of the line between possibility and plausibility of 'entitle[ment] to relief.'" *Id.* at \*9.

65 For discussion of this case, see Jeff Miles, "Analyzing the Federal Trade Commission's *North Texas Specialty Physicians Decision*," *The Health Lawyer* 18, no. 4 (April 2006), p. 1.

66 See, for example, Commissioner Jon Leibowitz, "Exclusion Payments to Settle Pharmaceutical Patent Cases: They're B-a-a-a-ck!" Second Annual In-House Counsel's Forum on Pharmaceutical Antitrust, Philadelphia, PA, April 24, 2006, available at FTC website.

67 *United States v. Microsoft Corp.*, 253 F.3d 34 (D.C. Cir. 2001).

68 Commission Decision of 24.03.2004 relating to a proceeding under Article 82 of the E.C. Treaty (Case COMP/C-3/37.792 Microsoft).

69 *In re Rambus, Inc.*, No. 9302 (F.T.C. July 31, 2006), available at <http://www.ftc.gov/os/adipiro/d9302/060802commissionopinion.pdf>, remedy ordered, *In re Rambus, Inc.*, No. 9302 (F.T.C.), available at <http://www.ftc.gov/os/adipiro/d9302/070205finalorder.pdf>.

70 *In re Union Oil Co. of Cal.*, No. 9305 (F.T.C. Mar. 4, 2003), available at <http://www.ftc.gov/os/2003/03/unocalcp.htm>, resolved by consent order, available at <http://www.ftc.gov/os/adipiro/d9305/050802do.htm>.

71 *Biovail Corp. and Elan Corp.*, File No. 011 0132 (F.T.C. June 27, 2002) (consent order), available at

cases that the defendant did not disclose a key patent to standards setting organization during the creation of a standard, and then “held up” the other members of the organization for high royalties.

In *Biovail*<sup>71</sup> and *Bristol-Myers Squibb*<sup>72</sup>, the FTC has challenge a pioneer firms listing of patents in the FDA’s “Orange Book” in order to block the entry of generic substitutes. Unlike the *Schering* type of FTC agreements cases between pioneer and generic drug manufacturers, the DOJ has not expressed opposition to challenging unilateral anticompetitive acts involving the alleged misuse of patents, as evidenced in a recent joint DOJ and FTC *amicus* brief submitted to the Second Circuit.<sup>73</sup>

It appears that at least the FTC will continue to pursue selective unilateral cases to the extent the courts will permit it. The agencies clearly have been giving a great deal of thought to the area, as illustrated by the resources the FTC and the DOJ have devoted to studying and holding hearings on the areas of patent protection in high tech industries and the pharmaceutical industry.<sup>74</sup> The agencies have also recently concluded a large number of hearings on single firm conduct, although much of the testimony has argued for restraint in pursuing such cases.<sup>75</sup> Senior officials have written articles and given speeches specifically addressing what they believe to be the appropriate approach to monopolization cases.<sup>76</sup> Senior officials at the DOJ and members of the Commission have also written articles and given a number of speeches on antitrust and intellectual property, devoting at least part of their analyses to potentially anticompetitive unilateral acts.<sup>77</sup>

Pure vertical restraints cases seem less likely to be pursued by either agency for several reasons. First, there does not appear to be a strong push for these cases by either agency’s senior officials. The FTC’s Commissioner Harbour gave a speech last year on vertical restraints, and concluded that “. . . I hope to see cutting-edge initiatives that clarify the law and impose appropriate remedies[.]”<sup>78</sup> However, to our knowledge neither the law nor the appropriate remedies have been substantially clarified to date – at least not in a way that would suggest more vertical restraints cases by the agencies.<sup>79</sup>

Second, the theoretical arguments put forward by the Chicago School of economics since the 1970s that vertical restraints are virtually always procompetitive, although questioned by many economists over the last decade, still appear to be heavily influencing both the agencies and the courts. There have been a great number of game theory based economics papers showing the potential for anticompetitive harm from vertical restraints, but there is relatively little empirical research showing whether such restraints damage competition.<sup>80</sup>

<http://www.ftc.gov/os/2002/06/biovailagreement.pdf>.

72 *Bristol-Myers Squibb Co.*, No. C-4076 (F.T.C. Apr. 18, 2003) (complaint), available at

<http://www.ftc.gov/os/2003/04/bristolmyerssquibbcmp.pdf>.

73 See Brief for the United States and Federal Trade Commission as Amici Curiae Supporting Plaintiffs-Appellants, Civ. Act. No. 06-5525-cv (S.D. N. Y.), May 30, 2007, available at <http://www.ftc.gov/os/2007/05/DDAVPCCommission-DoIBrief.pdf>. (“The dismissal of the *Walker Process* claim should not be affirmed on the basis that the plaintiffs lack antitrust standing as direct purchasers to bring such a claim.”)

74 See, for example, U.S. Department of Justice and the Federal Trade Commission, “Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition,” April 2007, available at

<http://www.ftc.gov/reports/innovation/P040101PromotingInnovationandCompetitionrpt0704.pdf>. This report discusses in detail the procompetitive aspects of intellectual property rights. In general, the report argues for a cautious approach to challenging potentially anticompetitive acts involving intellectual property, and as a matter of prosecutorial discretion the agencies will evaluate these issues using in effect full rule of reason analysis – including a balancing of anticompetitive effects and efficiencies.

75 See <http://www.ftc.gov/opa/2007/05/section2507.shtml>.

76 See, for example, DOJ AAG Thomas O. Barnett, “The Gales of Creative Destruction: The Need for Clear and Objective Standards for Enforcing Section 2 of the Sherman Act,” Opening Remarks for the Antitrust Division and Federal Trade Commission Hearings Regarding Section 2 of the Sherman Act, Washington DC, June 20 2006, available at <http://www.usdoj.gov/atrp/public/speeches/216738.htm>, and Kovacic (2007a).

77 See, for example, William Kovacic, FTC Commissioner, “The Importance of History to the Design of Competition Policy: The Federal Trade Commission and Intellectual Property,” 30 *Seattle University Law Review* 319 (2007). FTC Chairman Deborah Platt Majoras, “A Government Perspective on IP and Antitrust Law,” The IP Grab: The Struggle Between Intellectual Property Rights and Antitrust, the American Antitrust Institute, Washington DC, June 21 2006, available at <http://www.ftc.gov/speeches/majoras/060621aai-ip.pdf>; and Gerald F. Masouli, Deputy Assistant Attorney General at the Antitrust Division, “Intellectual Property and Competition: Four Principles for Encouraging Innovation,” Digital Americas 2006 Meeting, Intellectual Property and Innovation in the Digital World, Sao Paulo, Brazil, April 11, 2006, available at <http://www.usdoj.gov/atrp/public/speeches/215645.htm>.

78 Pamela Jones Harbour, “Vertical Restraints: Federal and State Enforcement of Vertical Issues,” ALI-ABA Course of Study, Product Distribution and Marketing, March 17-19, 2005, New Orleans, available at FTC web site.

79 Senior agency officials are also giving a great deal of thought to vertical cases. For example, see Kovacic (2007a) and J. Thomas Rosch, Commissioner of the FTC, “Vertical Restraints and Section 2, Washington D.C., June 13, 2007, available at

<http://www.ftc.gov/speeches/rosch/070613verticalrestraints.pdf>.

80 See, for example, Michael A. Salingor, Director of the FTC Bureau of Economics, “Looking for Keys under the Lamppost: Insights from Economics into Standards for Unilateral Conduct,” AB A Section of Antitrust Law, Economics and Section 2 Committees Brownbag, Washington, D.C., July 24, 2006, available at FTC web site; and Michael D. Whinston, *Lectures on Antitrust Economics* (MIT Press, 2006), at 3, pp. 133-197.

81 Michael A. Carrier, “Of *Trinko*, Tea Leaves, and Intellectual Property,” *The Journal of Corporation Law* 31, no. 2 (Winter 2006), pp. 357-373; Adam

Third, court decisions have made it more difficult for plaintiffs to win either monopolization or vertical restraints cases. The principal decision dealing with Section 2 at this time is *Verizon Communications v. Trinko*, 540 U.S. 398 (2004), which, largely in *dictum*, appears to close a lot of doors for Section 2 plaintiffs.<sup>81</sup> Refusals to deal that will be deemed anticompetitive now appear to be limited to the fact patterns of prior decisions, as *Trinko* evidenced a judicial reluctance to extend the doctrine. For a refusal to deal to be found predatory under Section 2, there will have to be a showing at least of an anticompetitive purpose and no legitimate business justification—a clear exclusion of rivals on a basis other than efficiency as the Supreme Court put it in *Aspen Skiing*.<sup>82</sup> Many scholars have suggested that the *Trinko* decision indicates a major retrenchment in antitrust law. As one commenter states “[The Court’s] thinly veiled swipes at antitrust – with its ‘considerable disadvantages,’ false positives, negative investment effects, and meddlesome courts – threaten to apply far beyond the facts in *Trinko*.” Another has written “The *Trinko* opinion could potentially immunize from antitrust scrutiny whole swathes of anticompetitive behavior.”<sup>84</sup> There are also questions about the continued viability of essential facility claims under Section 2 where there is no concerted action.

In its most recent term, in *Weyerhaeuser v. Ross-Simmons Hardwood*,<sup>85</sup> the Supreme Court extended the rule of *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*<sup>86</sup> to a claim for predatory buying, *i.e.*, allegedly cornering a market by buying up all available supply of a key input. In *Brooke Group*, the Court held low prices that are above cost ordinarily cannot constitute predatory conduct in violation of Sections 1 and 2 of the Sherman Act.<sup>87</sup> In *Weyerhaeuser*, the Court summarized its holding in *Brooke Group*, “Thus, we specifically decline to allow plaintiffs to recover for above-cost price cutting, concluding that ‘discouraging a price cut and . . . depriving consumers of the benefits of lower prices . . . does not constitute sound antitrust policy.’”<sup>88</sup> The Court held in *Weyerhaeuser*, “Consequently, only higher bidding that leads to below-cost pricing in the relevant output market will suffice as a basis for liability for predatory bidding.”<sup>89</sup> The Supreme Court’s decision in *Weyerhaeuser* could have a major impact on Section 2 jurisprudence, possibly further narrowing the scope of predatory conduct.

In the meantime, there are a number of Court of Appeals decisions that provide some assistance in defining the scope of exclusionary conduct, including *United States v. Microsoft*, (tying and exclusive dealing);<sup>90</sup> *Conwood Co. v. United States Tobacco Co.*, (dirty tricks, *i.e.*, removing rival’s display racks);<sup>91</sup> and *LePage’s v. 3M*, (bundled rebates).<sup>92</sup> Given *Trinko* and *Weyerhaeuser*, it is unclear what force any of these earlier decisions will continue to have. The law seems to have moved a long way from the rather simply expressed and easily workable definition of predatory conduct in *Aspen Skiing*, excluding rivals on a basis other than efficiency. Some lower court decisions, such as Judge Easterbrook’s decision on leveraging in *Schor v. Abbott*, \_\_\_ F.3d \_\_\_, 2006-2 CCH Trade Cas. Paragraph 75,354 (7th Cir. July 26, 2006), also suggest a clear narrowing of what remains of monopolization in the court’s eyes.<sup>93</sup>

As to vertical restraints, in its latest term, the Supreme Court overruled *Dr. Miles Medical Co. v. John D. Park & Sons Co.*,<sup>94</sup> and held that minimum resale price maintenance would no longer be *per se* illegal, but would be subject to the rule of reason.<sup>95</sup> This change will place a significantly greater burden of proof on plaintiffs.

Candeub, “*Trinko* and Re-Grounding the Refusal to Deal Doctrine,” *University of Pittsburgh Law Review* 66, no. 4 (Summer 2005), pp. 821-870.

82 *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*, 472 U.S. 585 (1985).

83 Michael A. Carrier, “Of *Trinko*, Tea Leaves, and Intellectual Property,” *Journal of Comparative Law* 357, 373 (Winter 2006).

84 Adam Candeub, “*Trinko* and Re-Grounding the Refusal to Deal Doctrine,” 66 *University of Pittsburgh Law Review* 821, 823 (2005).

85 \_\_\_ U.S. \_\_\_, 127 S.Ct. 1069 (2007).

86 509 U.S. 209 (1993).

87 The Court said in *Brooke Group*, “[W]e have rejected elsewhere the notion that above-cost prices that are below general market levels or the costs of a firm’s competitors inflict injury to competition cognizable under the antitrust laws.” 509 U.S. at 223. In *Brooke Group*, the Court also quoted its language in *Atlantic Richfield Co. v. USA Petroleum Co.*, (495 U.S. 328, 334-341 (1990)) “Low prices benefit consumers regardless of how those prices are set, and so long as they are above predatory levels, they do not threaten competition . . . . We have adhered to this principle regardless of the type of antitrust claim involved.” *Id.*, quoting *ARCO*, 498 U.S. at 340.

88 127 S.Ct. at 1074-75.

89 *Id.*, at 1078.

90 253 F.3d 34 (D.C. Cir. 2001).

91 290 F.3d 768 (6th Cir. 2002).

92 324 F.3d 141 (3d Cir. 2003).

93 See, also, *Smith Wholesale Co., Inc. v. R.J. Reynolds Tobacco Co.*, 477 F.3d 854 (6th Cir. 2007); *Concord Boat Corp. v. Brunswick Corp.*, 207 F.3d 1039 (8th Cir.), cert. denied, 531 U.S. 979 (2000). These decisions also cast doubt on the continued viability of *LePage’s*.

94 220 U.S. 373 (1911).

95 *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, \_\_\_ S.Ct. \_\_\_, 2007 WL 1835892 (June 28, 2007).

96 *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004); *Volvo Trucks No. Am., Inc. v. Reeder-Simco GMC, Inc.*,

Finally, there is another trend in the case law that cannot be ignored in forecasting the likely future of antitrust enforcement, at least during the remainder of the Bush administration. The present Supreme Court is active in antitrust in a predictable and interesting manner, as has been the participation of Solicitor General. Beginning with *Trinko*, the Court has granted certiorari in eight antitrust cases, most of these involving some form of alleged unilateral anticompetitive behavior.<sup>96</sup> In each case, the plaintiff prevailed in the Court of Appeals. In each case, the Government took the side of the defendant in the Supreme Court, and argued for a ruling that would make antitrust enforcement more restricted and difficult. In each case, the Supreme Court ruled for the defendant, even going beyond the relief sought by the Government in *Credit Suisse v. Billing* to provide blanket immunity for the conduct at issue. In each case, except *Trinko*, which is replete with *dicta* making private enforcement more difficult, the Court's rulings were narrow and limited, providing relatively little guidance beyond the holding of the decision. Thus, the enforcement agencies have succeeded repeatedly before the Supreme Court in limiting to some degree the scope of their own enforcement powers. Based on these filings, the agencies as currently constituted do not appear to envision or desire for themselves a particularly expansionist role in the enforcement of the antitrust laws.

Accordingly, it seems that absent more empirical research showing anticompetitive effects from monopolistic practices and vertical restraint, substantial new agency initiatives, and more plaintiff favorable court rulings, the agencies are likely only pursue these types of cases in limited circumstances. Given recent FTC enforcement actions, such cases are most likely to involve standards setting organizations or pioneer drug manufacturers use of the Orange Book to delay generic entry.

## V. OTHER AREAS

Certain types of cases have been abandoned for years by the agencies, and show no sign of returning. For example, up until the 1970s the FTC frequently brought price discrimination cases under the Robinson-Patman Act. The FTC has investigated one or two of these cases in the 1980s and 1990s, and we see no evidence of any renewed enforcement here. Moreover, court decisions have made these types of cases more difficult to win for the plaintiffs.<sup>97</sup>

There have been other types of cases that are not part of the mainstream classifications of mergers, agreements among competitors, and monopolization that were pursued in the 1980s and 1990s, and there is evidence that the agencies might bring some of these types of cases in the future. The FTC brought "facilitating practices" and "invitations to collude" cases in the 1980s and 1990s, but there has been relatively little activity in these areas in a number of years.<sup>98</sup> The FTC brought these cases at least in part under what it considered to be a violation of Section 5 of the FTC Act's unfair method of competition, rather than under the traditional antitrust Sherman and Clayton Acts. However, in March 2006 the FTC challenged and obtained a consent by a unanimous vote in the *Valassis* invitation to collude case based solely on Section 5.<sup>99</sup> One case does not necessarily make a trend for future enforcement, but at least one Commissioner has stated that he believes there are unfair methods of competition cases that can be challenged under the FTC's Section 5 and would not fit into the practices challenged under the traditional antitrust statutes.<sup>100</sup> The unanimous vote in that case suggests there may be renewed enforcement in this area.

## VI. CONCLUSION

We have analyzed recent past agencies enforcement in several areas, and what factors appear to have been and will be influencing it. Based on this analysis, we anticipate that federal challenges to

546 U.S. 164 (2006); *Illinois Tool Works, Inc. v. Independent Ink, Inc.*, 547 U.S. 28 (2006); *Texaco, Inc. v. Dagher*, 547 U.S. 1 (2006); *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc.*, \_\_\_ U.S. \_\_\_, 127 S.Ct. 1069 (2007); *Bell Atlantic Corp. v. Twombly*, \_\_\_ U.S. \_\_\_, 127 S.Ct. 1955, 2007 WL 1461066 (May 21, 2007); *Leggin Creative Leather Prods., Inc. v. PSKS, Inc.*, \_\_\_ U.S. \_\_\_, \_\_\_ S.Ct. \_\_\_ (2007) WL 1835892 (June 28, 2007); *Billing v. Credit Suisse First Boston, Ltd.*, \_\_\_ U.S. \_\_\_, \_\_\_ S.Ct. \_\_\_ 2007 WL 1730141 (June 18, 2007).

<sup>97</sup> See Margaret Zwisler, "Volvo Trucks v. Reeder-Simco: Judicial Activism at the Supreme Court?" *Antitrust* 20, no. 3 (Summer 2006), p. 40.

<sup>98</sup> See Veronica Kayne, "Section 5 of the FTC Act: Not All Gaps Need Filling," *Antitrust Bulletin* 49, no. 3 (Fall 2004), p. 783.

<sup>99</sup> *In the matter of Valassis Communications, Inc.* FTC File No. 051 0008 (March 16, 2006) available at <http://www.ftc.gov/os/caselist/0510008/0510008.htm>

<sup>100</sup> J. Thomas Rosch, "Perspectives on Three Recent Votes: The Closing of the Adelphia Communications Investigation, the Issuance of the Valassis Complaint & the Weyerhaeuser Amicus Brief," before The National Economic Research Associates 2006 Antitrust & Trade Regulation Seminar, Santa Fe, New Mexico, July 6, 2006 available at <http://www.ftc.gov/speeches/rosch/Rosch-NERA-Speech-July6-2006.pdf>.

mergers are likely to continue, but at the current relatively low rate. Merger enforcement actions have been fewer at least since 2004, even taking into account the reduced level of merger activity. Recent court decisions are likely to discourage or prevent more activism in challenging mergers, even with a change in the administration.

The DOJ will continue to be active in investigating criminal price fixing and market allocation cases, but the number of challenges will likely remain relatively low absent a change in policy. The FTC has continued to challenge non-criminal agreements among competitors' cases, recently using its "quick look" approach. However, that approach has not been completely embraced by the courts, and might be limited by decisions such as *Dagher*.

Section 2 is not a likely area of much federal enforcement in the near future, although the FTC may continue to challenge company actions similar to those in its recent standards and Orange Book listings cases. This low level of activity is due to both existing agency policies and the courts. Other areas are not likely to be revived in the foreseeable future, except perhaps in the area of FTC challenges based primarily on its Section 5 powers.

In general, we do not anticipate significantly more aggressive antitrust enforcement from the agencies in the near future. Even if there were a major shift in antitrust policy toward more actively challenging potentially anticompetitive acts, the courts have placed and likely would place more limits on what the agencies can do.

# THE ANCILLARY RESTRAINTS DOCTRINE AFTER *DAGHER*

Gregory J. Werden\*

U.S. Department of Justice, Antitrust Division  
Washington, DC

The common-law concept of an “ancillary restraint” was imported into Sherman Act jurisprudence by Judge Taft’s 1898 *Addyston Pipe* opinion,<sup>1</sup> an important early attempt to come to grips with the meaning of section 1 of the Sherman Act.<sup>2</sup> Under the common law<sup>3</sup> Judge Taft concluded, “no conventional restraint of trade can be enforced unless the covenant embodying it is merely ancillary to the main purpose of a lawful contract, and necessary to protect the covenantee in the full enjoyment of the legitimate fruits of the contract, or to protect him from the dangers of an unjust use of those fruits by the other party.”<sup>4</sup>

Judge Taft’s innovation went largely unnoticed until highlighted and extended in 1959 by Robert Bork,<sup>5</sup> and the most prominent modern statement of the ancillary restraints doctrine may be that in Bork’s 1986 opinion for the D.C. Circuit in *Rothery*: “To be ancillary, and hence exempt from the per se rule, an agreement eliminating competition must be subordinate and collateral to a separate, legitimate transaction. The ancillary restraint is subordinate and collateral in the sense that it serves to make the main transaction more effective in accomplishing its purpose.”<sup>6</sup>

The Supreme Court touched on the ancillary restraints doctrine several times. Justice Stewart’s concurring and dissenting opinion in *Schwinn* noted that the common law permitted restraints “when ancillary to a legitimate business purpose and not unduly anticompetitive” and cited Judge Taft’s opinion for the proposition that the “doctrine of ancillary restraints was assimilated into the jurisprudence of this country in the nineteenth century.”<sup>7</sup> Justice Stevens’ opinion for the Court in *Professional Engineers* cited Judge Taft’s opinion for the proposition that the rule of reason is the “standard for testing the enforceability of covenants in restraint of trade which are ancillary to a legitimate transaction.”<sup>8</sup> Justice Stevens’ dissenting opinion in *Business Electronics* discussed the ancillary restraints doctrine at length, remarking that Judge Taft’s “opinion is universally accepted as authoritative.”<sup>9</sup> Finally, Justice Thomas’ opinion for the Court in *Dagher* explained that the ancillary restraints doctrine “governs the validity of restrictions imposed by a legitimate business collaboration,

\* Senior Economic Counsel, Antitrust Division, U.S. Department of Justice. The views expressed herein are not purported to represent those of the U.S. Department of Justice.

1 *United States v. Addyston Pipe & Steel Co.*, 85 F. 271 (6th Cir. 1898), *aff’d as modified*, 175 U.S. 211 (1899).

2 Contemporaneous were Justice Peckham’s landmark opinions in *United States v. Joint Traffic Ass’n*, 171 U.S. 505 (1898); *United States v. Trans-Missouri Freight Ass’n*, 166 U.S. 290 (1897). Justice Peckham also authored the opinion affirming of Judge Taft’s decision.

3 The “classic ancillary” restraint is an agreement by the seller of a business not to compete within the market.” *Bus. Elecs. Corp. v. Sharp Elecs. Corp.*, 485 U.S. 717, 729 n.3 (1988) (citing *Mitchel v. Reynolds*, (1711) 24 Eng. Rep. 247 (K.B.)). The classic ancillary restraint has been the subject of some antitrust cases. *E.g.*, *Lektro-Vend Corp. v. Vendo Co.*, 660 F.2d 255, 264-69 (7th Cir. 1981) (rejecting Section 1 challenge to a covenant not to compete associated with an acquisition).

4 *Addyston Pipe*, 85 F. at 282. Judge Taft’s opinion is often cited for his colorful rejection of consideration of the reasonableness of non-ancillary restraints. *Id.* at 283-84 (“It is true that there are some cases in which the courts, mistaking, as we conceive, the proper limits of the relaxation of the rules for determining the unreasonableness of restraints of trade, have set sail on a sea of doubt, and have assumed the power to say, in respect to contracts which have no other purpose and no other consideration on either side than the mutual restraint of the parties, how much restraint of competition is in the public interest, and how much is not. The manifest danger in the administration of justice according to so shifting, vague, and indeterminate a standard would seem to be a strong reason against adopting it.”).

5 Robert H. Bork, *Ancillary Restraints and the Sherman Act*, 15 A.B.A. ANTITRUST SEC. PROC. 211 (1959). *See also* ROBERT H. BORK, *THE ANTITRUST PARADOX* 26-30, 270-79 (rev. ed. 1993) (1978); Robert H. Bork, *The Rule of Reason and the Per Se Concept: Price Fixing and Market Division*, 74 *YALE L.J.* 775, 797-98 (1965).

6 *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.*, 792 F.2d 210, 224 (D.C. Cir. 1986).

7 *United States v. Arnold, Schwinn & Co.*, 388 U.S. 365, 392 (1967).

8 *Nat’l Soc’y of Prof’l Eng’rs v. United States*, 435 U.S. 679, 689 (1978).

9 *Bus. Elecs. Corp. v. Sharp Elecs. Corp.*, 485 U.S. 717, 737-39 (1988). Judge Taft’s opinion may be considered authoritative because he became Chief Justice and because joining in the opinion were Justice Harlan and future Justice Lorton.

such as a business association or joint venture, on nonventure activities” and that the doctrine “has no application” when “the business practice being challenged involves the core activity of the joint venture itself.”<sup>10</sup>

This article addresses the application of the ancillary restraints doctrine to restraints associated with joint ventures formed by competitors.<sup>11</sup> In particular, five specific questions are addressed: First, exactly what is the “ancillary restraints doctrine;” what is at stake in deciding whether a restraint is ancillary? Second, what restraints associated with a joint venture are separate from the venture but collateral to it and thus the subject of the ancillary restraints doctrine, as opposed to what the *Dagher* Court referred to as “core activity of the joint venture itself”? Third, what makes a collateral restraint “ancillary;” what nexus to the venture and the accomplishment of its procompetitive purposes must the restraint have to be deemed “ancillary”? Fourth, precisely what is the “test” for ancillarity, and how are burdens allocated in litigation? Finally, what should we make of *Topco* and *Citizen Publishing*,<sup>12</sup> two cases in which the Supreme Court applied the *per se* rule to joint venture restraints?

## I. WHAT IS THE “ANCILLARY RESTRAINTS DOCTRINE”?

Judge Taft took the view that ancillary restraints were lawful and non-ancillary restraints were unlawful. Were that true today, the ancillary restraints doctrine would be the keystone of Section 1 analysis, but surely that is not the role the doctrine plays. Robert Bork insightfully explained that the “common law doctrine of ancillary restraints offers the Sherman Act not content but form” and that “the sole function of the concept of ancillarity under the Sherman Act should be to point out instances when *per se* illegality should not attach and to confine the exceptions to their proper scope.”<sup>13</sup> Citing Professor Bork, the Ninth Circuit held: “The proper function of ancillarity in antitrust analysis ‘is to remove [in some instances] the *per se* label from restraints otherwise falling within the category.’”<sup>14</sup>

Professor Hovenkamp indicates that “once a restraint is found to be ancillary, the court pursues its inquiry with a presumption of lawfulness and requires . . . proof of power and effects,” and he suggests that an ancillary restraint is unlawful if it has “output-reducing tendencies (anticompetitive effects) that are not offset by reasonable justifications or defense.”<sup>15</sup> This treatment presupposes that a restraint is declared ancillary on the basis of a mere potential to facilitate the accomplishment of a joint venture’s legitimate objectives.<sup>16</sup> In that event, the “reasonableness” issue is clearly distinct from the “ancillarity” issue, as in Judge Easterbrook’s articulation of the ancillary restraints doctrine in *Polk Bros.*:

A court must ask whether an agreement promoted enterprise and productivity at the time it was adopted. If it arguably did, then the a court must apply the Rule of Reason to make a more discriminating assessment. . . .

If the restraint, viewed at the time it was adopted, may promote the success of the more extensive cooperation, then the court must scrutinize things carefully under the Rule of Reason.<sup>17</sup>

An alternative treatment merges the test for ancillary with the “reasonableness” issue. Although perhaps never explicitly stated, some cases appear to have in mind a concept of ancillarity

10 *Texaco Inc. v. Dagher*, 547 U.S. 1, 7-8 (2006).

11 In his 1959 article, Bork remarked that “the most fruitful area for the application of the common law of ancillary restraints to the Sherman Act is in the analysis of restraints ancillary to joint ventures.” Bork, *Ancillary Restraints*, *supra* note 5, at 224. Professor Hovenkamp has gone so far as to suggest that “the most frequent application of the rule of reason involves restraints that are ‘ancillary’ to some underlying productive joint venture.”

12 *United States v. Topco Assocs.*, 405 U.S. 596 (1972); *Citizen Publ’g Co. v. United States*, 394 U.S. 131 (1969).

13 Bork, *Ancillary Restraints*, *supra* note 5, at 216, 227. See also *id.* at 226 (“It seems best . . . to restrict the concept of ancillarity to those horizontal restraints – agreements to divide markets and to fix prices – which are usually thought of as illegal *per se*.”).

14 *Aydin Corp. v. Loral Corp.*, 718 F.2d 897, 901 (9th Cir. 1983) (quoting Bork, *Ancillary Restraints*, *supra* note 5, at 212). See also *Los Angeles Mem’l Coliseum Comm’n v. NFL*, 726 F.2d 1381, 1395 (9th Cir. 1984) (the practical “effect of a finding of ancillarity is to ‘remove the *per se* label from restraints otherwise falling within that category’”) (quoting Bork, *Ancillary Restraints*, *supra* note 5, at 212).

15 11 HOVENKAMP, *supra* note 11, Paragraphs 1904, 1912c4, at 227, 326.

16 Although not entirely clear, that appears to be Professor Hovenkamp’s suggested approach. Most notably, in defining an ancillary restraint (see *infra* text at note 49), he uses the phrase “at least upon initial examination.” *Id.* Paragraph 1912c, at 320.

17 *Polk Bros. v. Forest City Enters.*, 776 F.2d 185, 189 (7th Cir. 1985).

more like that of Judge Taft, effectively making it coextensive with the application of the rule of reason: Only reasonable restraints are considered ancillary, so all ancillary restraints are lawful.

I favor yet a third treatment in which the ancillary issue determines whether a joint venture's collateral restraints are assessed separate and apart from the joint venture. Non-ancillary restraints are assessed separately, which may result in per se condemnation. Ancillary restraints are assessed as together with the formation of the joint venture – the whole package being evaluated under the rule of reason if the joint venture itself is legitimate.<sup>18</sup> An ancillary restraint may alter the balance of competitive costs and benefits from the formation of the joint venture and could tip the rule of reason scales against the venture.<sup>19</sup>

Roughly the same idea was expressed by Chief Judge Ginsburg in his *PolyGram* opinion.<sup>20</sup> In analyzing a hypothetical involving the introduction of a new product by a joint venture of major automobile producers, he observed that, “if the only way a new product can be profitably introduced is to restrain the legitimate competition of older products, then one must seriously wonder whether consumers are genuinely benefited by the new product.”<sup>21</sup> Although Judge Ginsburg did not explain his observation in terms of antitrust doctrine, the proper analysis of his hypothetical could be that the restraint is ancillary and causes the formation of the joint venture to be unlawful under the rule of reason.

In sum, the “ancillary restraints doctrine” is understood in this article to hold that restraints collateral to the formation of a joint venture are assessed together with the formation of the joint venture if ancillary, but assessed separately from the formation of the joint venture if non-ancillary. For price or output restraints collateral to a legitimate joint venture, one implication is that ancillary restraints are subject to the rule of reason and non-ancillary restraints may be per se illegal.

## II. WHAT IS A “COLLATERAL” RESTRAINT?

As Judge Taft interpreted the common law, “Restrictions in the articles of partnership upon the business activity of the members, with a view of securing their entire effort in the common enterprise, were, of course, only ancillary to the main end of the union, and were to be encouraged.”<sup>22</sup> Adopting essential elements of the common law, Judge Bork's *Rothery* opinion states that a restraint can be ancillary only if “subordinate and collateral to a separate, legitimate transaction.”<sup>23</sup> Judge Easterbrook similarly remarked in *Polk Bros.* that ancillary restraints “are part of a larger endeavor whose success they promote.”<sup>24</sup> Furthermore, the Eleventh Circuit recently held that: “In order for a condition to be ancillary, an agreement limiting competition must be secondary and collateral to an independent and legitimate transaction.”<sup>25</sup>

For charge card joint ventures, agreements that members not issue certain competing cards were collateral restraints.<sup>26</sup> In an agreement dissolving a law partnership, a territorial restriction on advertising by the former partners was a collateral restraint.<sup>27</sup> In a joint venture to build and operate a building containing two separate retail stores, a restraint on which products could be sold was

18 See Gregory J. Werden, *Antitrust Analysis of Joint Ventures: An Overview*, 66 ANTITRUST L.J. 701, 706-07, 734 (1998).

19 The 1995 Antitrust Guidelines for the Licensing of Intellectual Property issued jointly by the U.S. Department of Justice and the FTC take a similar approach:

If the Agencies conclude that the restraint has, or is likely to have, an anticompetitive effect, they will consider whether it is reasonably necessary to achieve procompetitive efficiencies. If the restraint is reasonably necessary, the Agencies will balance the procompetitive efficiencies and the anticompetitive effects to determine the probable net effect on competition in each relevant market.

Section 4.2, reprinted in 4 Trade Reg. Rep. (CCH) Paragraph 13,132. Such an analysis also appears to have been suggested by Bork, *Ancillary Restraints*, *supra* note 5, at 228.

20 *PolyGram Holding, Inc. v. FTC*, 416 F.3d 29 (D.C. Cir. 2005).

21 *Id.* at 38.

22 *United States v. Addyston Pipe & Steel Co.*, 85 F.271, 280 (6th Cir. 1898), *aff'd as modified*, 175 U.S. 211 (1899).

23 *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.*, 792 F.2d 210, 224 (D.C. Cir. 1986). In his 1959 article, Bork cited two primary examples of restraints that are collateral and potentially ancillary: “The agreement of the seller of a business not to compete with the business sold; and . . . the agreement of the participant in a joint venture not to compete with that venture.” Bork, *Ancillary Restraints*, *supra* note 5, at 211.

24 *Polk Bros. v. Forest City Enters.*, 776 F.2d 185, 189 (7th Cir. 1985).

25 *Schering-Plough Corp. v. FTC*, 402 F.3d 1056, 1073 (11th Cir. 2005), *cert. denied*, 126 S. Ct. 2929 (2006). See also 11 HOVENKAMP, *supra* note 11, Paragraph 1912c2, at 322 (“when a restraint is challenged, one generally looks for some ‘underlying’ agreement or arrangement among the parties that has at least the potential to reduce costs, provide a new or improved product, or otherwise increase output”).

26 *United States v. Visa U.S.A., Inc.*, 344 F.3d 229, 237, 243 (2d Cir. 2003), *cert. denied*, 543 U.S. 811 (2004).

27 *Blackburn v. Sweeney*, 53 F.3d 825 (7th Cir. 1995).



collateral.<sup>28</sup> And a territorial restriction was collateral to an arrangement in which truck leasing companies provided service for each others' trucks.<sup>29</sup> Courts also have treated as collateral some restraints on membership in joint ventures.<sup>30</sup>

In *Dagher*, the Ninth Circuit applied the ancillary restraints doctrine to a restraint imposed on the actions of a joint venture.<sup>31</sup> The joint venture combined the domestic petroleum refining and gasoline marketing operations of Texaco and Shell, entirely ending competition between them in the affected markets.<sup>32</sup> The separate Texaco and Shell brands of gasolines were sold by the joint venture subject to the restraint that they be priced the same,<sup>33</sup> and this restraint was challenged as per se illegal price fixing.<sup>34</sup> The district court granted defendants' motion for summary judgment, but the Ninth Circuit reversed.

The Ninth Circuit observed that the issue was "whether the price fixing is 'naked' (in which case the restraint is illegal) or 'ancillary' (in which case it is not)."<sup>35</sup> Invoking the ancillary restraints doctrine, the court held that the key question was whether the restraint was "reasonably necessary to further the legitimate aims of the joint venture."<sup>36</sup> Because the defendants had "failed to offer any explanation of how their unified pricing of the distinct Texaco and Shell brands of gasoline served to further the ventures' legitimate efforts to produce better products or capitalize on efficiencies," the Ninth Circuit held that the restraint might well be per se unlawful.<sup>37</sup>

The Supreme Court took a quite different view than the Ninth Circuit, holding that, "because Texaco and Shell Oil did not compete with one another in the relevant market," the restraint was not "price fixing in the antitrust sense."<sup>38</sup> The Court further held that the Ninth Circuit erred in invoking the ancillary restraints doctrine, which "has no application" when "the business practice being challenged involves the core activity of the joint venture itself" such as "pricing of the very goods sold by" the venture.<sup>39</sup> Rather, the Court indicated that the ancillary restraints doctrine "governs the validity of restrictions imposed by a legitimate business collaboration, such as a business association or joint venture, on nonventure activities."<sup>40</sup>

The Supreme Court's analysis recognizes that constraints joint venture participants place on the conduct of the venture itself, within the ambit of its operation, are quite different than constraints they place on their own conduct outside the venture. Participants in a manufacturing joint venture almost certainly determine, and hence limit, the products the venture produces, and they may agree to limit the geographic area in which the venture operates, the customers to which the venture sells, or the prices it charges. Such constraints are not collateral restraints; rather, they are integral parts of the joint venture that cannot be separated from it.

Although the Court left participants in a joint venture entirely free to decide what their joint venture does and how it is done, that in no way suggests that joint ventures escape meaningful antitrust scrutiny.<sup>41</sup> The formation of a joint venture may be unlawful because of decisions its

28 *Polk Bros. v. Forest City Enters.*, 776 F.2d 185 (7th Cir. 1985). Judge Easterbrook specifically rejected the district court's determination that the restraint was not collateral but rather was "an integral part of" the joint venture. *Id.* at 190.

29 *Gen. Leaseways, Inc. v. Nat'l Truck Leasing Ass'n*, 744 F.2d 588 (7th Cir. 1984).

30 *See, e.g., SCFC ILC, Inc. v. Visa USA, Inc.*, 36 F.3d 958 (10th Cir. 1994) (agreement among members of a charge card joint venture to restrict membership); *Sullivan v. NFL*, 34 F.3d 1091 (1st Cir. 1994) (agreement among members of a sports league to ban public ownership of teams).

31 *Dagher v. Saudi Refining Inc.*, 369 F.3d 1108 (9th Cir. 2004), *rev'd*, 547 U.S. 1 (2006).

32 *Id.* at 1111-12.

33 *Id.* at 1112.

34 Plaintiffs also argued that the restraint was unlawful under the "quick look" if not per se illegal. Finding a possible per se violation, the Ninth Circuit did not address the alternative "quick look" argument. *Id.* at 1116 n.7. The Supreme Court rejected liability under the "quick look" "for the same reasons that [it found] per se liability is unwarranted." 547 U.S. at 7 n.3.

35 *Dagher*, 369 F.3d at 1118.

36 *Id.* at 1121.

37 *Id.* at 1122. The court found the record "close to establishing that the price-fixing scheme was sufficiently unrelated to accomplishing the legitimate objectives of the joint venture as to justify granting the plaintiffs' motion for summary judgment" but concluded that denying both sides' summary judgment motions was proper under the circumstances. *Id.* at 1122 n.16.

38 547 U.S. at 5-6.

39 *Id.* at 7-8.

40 *Id.* at 7. According to Professor Hovenkamp, "the Court appeared to leave a large area of per se liability for restraints that (1) operate on the participants' intra-venture activity but (2) are not part of the core," and he argues that many such restraints should be viewed as ancillary and therefore not subject to the per se rule. PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW Paragraph 2132, at 509-10 (Supp. 2006). A better reading of the decision, however, is that no constraints on intra-venture activity are assessed separate from the venture itself and hence none are subject to the per se rule if the joint venture is legitimate.

41 Professor Hovenkamp worries that the Court's "core" concept could allow a joint venture to "serve as a front for anticompetitive collusion." *Id.* Paragraph 2132, at 508. But a sham joint venture serving as a front for collusion is per se illegal, and a legitimate joint venture primarily orchestrating a cartel can be found unlawful under a highly abbreviated application of the rule of reason.

participants have made about what it does or how it does it. But if the formation of a venture enhances competition despite the constraints placed upon its operations, the venture is not vulnerable to attack on the grounds that the constraints constitute free-standing violations of Section 1.

Chief Judge Ginsburg's *PolyGram* decision<sup>42</sup> does not mention the ancillary restraints doctrine but has been read to restrict its application. The case concerned a joint venture to produce and market the album of the third Three Tenors concert, and the participants in the venture were the two firms that had separately produced and marketed the albums of the Three Tenors' first two concerts.<sup>43</sup> The FTC held that the firms violated the antitrust laws by agreeing not to promote the first two Three Tenors albums when they introduced the third, and the D.C. Circuit denied the petition for review of the FTC's decision. The agreement was treated by the FTC and the court as a restraint on the actions of the joint venture participants outside the scope of the joint venture, and both rejected as a matter of law the justification that the agreement was a permissible effort to prevent free riding.<sup>44</sup> Commentators have argued that the court thereby precluded treating as ancillary any restraints on competition outside the venture.<sup>45</sup> It is doubtful, however, that Chief Judge Ginsburg overruled *sub silentio* the D.C. Circuit's controlling precedent (*Rothery*), and *Dagher* has now made clear that ancillary restraints doctrine "governs the validity of restrictions . . . on nonventure activities."<sup>46</sup>

In sum, "ancillary restraints doctrine" applies only to "collateral" restraints associated with joint ventures. Such restraints regulate not what the joint venture does or how it is done, but rather the conduct of the joint venture's participants outside the venture.

### III. WHAT MAKES A COLLATERAL RESTRAINT "ANCILLARY"?

The Supreme Court has noted that an ancillary restraint in the common law "enhances the value of the contract, or permits the 'enjoyment of [its] fruits.'"<sup>47</sup> Robert Bork explained that, in the context of joint ventures, an "agreement not to compete may be necessary in order that each participant may be safe in contributing his best efforts and his resources to the joint enterprise."<sup>48</sup> And Professor Hovenkamp has indicated that, as a general matter:

An ancillary restraint is one that is reasonably related to a joint venture or transaction that, at least upon initial examination, promises to increase output, reduce costs, improve product quality, or otherwise benefit consumers. As a result, the profitability of an ancillary restraint depends not on the market power of its participants but on their ability to reduce their costs or increase demand for their product.<sup>49</sup>

A collateral restraint may make a joint venture operate more efficiently: A requirement that participants in a manufacturing joint venture buy exclusively from the venture could facilitate the venture's realization of economies of scale.

A collateral restraint may prevent a joint venture participant from appropriating an undue share of the venture's benefits: Exclusive distribution territories for a brand created and promoted by a joint venture could ensure that each participant in the joint venture reaps only its intended reward.

A collateral restraint may prevent non-participants from appropriating joint venture benefits for which they have not shared the costs: A ban on sub-licensing a research joint venture's intellectual property to non-participants could ensure that firms not sharing in the costs and risks of the venture do not share in its advantages.

42 *PolyGram Holding, Inc. v. FTC*, 416 F.3d 29, 31-32 (D.C. Cir. 2005).

43 *Id.* at 31-32.

44 *Id.* at 33, 37-38.

45 See David L. Meyer & Derek Ludwin, *Three Tenors and the Section 1 Analytical Framework: A Continuum Drawn with Bright Lines*, ANTITRUST, Fall 2005, at 63, 68 (citing 416 F.3d at 38).

46 *Dagher*, 547 U.S. at 7.

47 *Bus. Elecs. Corp. v. Sharp Elecs. Corp.*, 485 U.S. 717, 729 n.3 (1988) (quoting *United States v. Addyston Pipe & Steel Co.*, 85 F. 271, 282 (6th Cir. 1898)).

48 Bork, *Ancillary Restraints*, *supra* note 5, at 217-18; see also *id.* at 228-29 (illustrating the point).

49 11 HOVENKAMP, *supra* note 11, Paragraph 1912c, at 320; see also *id.*, Paragraph 1904, at 227 ("To say that a restraint is 'ancillary' is to conclude that it is an essential or at least an important part of some arrangement that has potentially redeeming values.").

A collateral restraint also could prevent unintended competitive consequences that might make the venture uneconomic. Consider a joint venture that supplies its two participants with an industrial commodity each uses to manufacture a product not in competition with the other's product. By lowering their production costs, the formation of the venture could induce each participant to begin making the other's product. This effect could be forestalled by restraints of various sorts.

Whether a collateral restraint is ancillary depends on the particular circumstances faced by the participants in the joint venture. Consider two distribution joint ventures in which participants in joint ventures adopt restraints eliminating price competition. First, suppose joint venture participants sell competing differentiated consumer products marketed primarily through the Internet, and that the venture owns and operates facilities for taking, processing, and fulfilling orders. An agreement between the participants to sell their competing products at the same price is most unlikely to be ancillary because the benefits of sharing on-line sales functions almost certainly do not depend on collective pricing. Second, suppose that a joint venture builds and operates a facility from which its participants dispense a liquid chemical to their customers. An agreement between the participants to sell their competing products at the same price may be ancillary because they may not be able to sustain two different prices.

In *Polk Bros.*, a restraint prohibiting each of the two stores sharing a common site from selling certain products sold by the other was found to be ancillary and lawful.<sup>50</sup> The site was developed to provide premises for two retailers selling complementary lines of products. Polk Bros. sold appliances and home furnishings, and Forest City sold tools and materials for home improvement. At the outset, Forest City agreed not to sell appliances except for built-in models, and Polk Bros. agreed not to sell many enumerated products including Toro and Lawn Boy lawn mowers.<sup>51</sup> Both parties apparently breached their obligations, and when Polk sought an injunction against Forest City, the district court declared the covenant not to compete to be a per se violation of Section 1. The Seventh Circuit reversed, finding that "the restrictive covenant made cooperation possible" which in turn "increased the amount of retail space available and was at least potentially beneficial to consumers."<sup>52</sup> Judge Easterbrook's opinion explained that a "restraint is ancillary when it may contribute to the success of a cooperative venture that promises greater productivity and output" and that the specific question for a court is "whether an agreement promoted enterprise and productivity at the time it was adopted."<sup>53</sup>

It is easy to construct scenarios in which collateral restraints are not ancillary. The Ninth Circuit's *Dagher* opinion offers the example of a research joint venture formed by two soft drink companies to create a new flavor that is accompanied by an agreement between the companies to fix the price of products unrelated to the venture.<sup>54</sup> In this hypothetical, as in many others, "the organic connection between the restraint and the cooperative needs of the enterprise that would allow us to call the restraint a merely ancillary one is missing."<sup>55</sup> A corollary of the necessity of an "organic connection" is that a "restraint cannot be justified solely on the ground that it increases the profitability of the enterprise."<sup>56</sup> In the soft drink example, the price fixing in unrelated products may make the joint venture profitable when it otherwise would not be, but that does not make the price fixing ancillary.

In sum, a restraint collateral to the formation of a legitimate joint venture is "ancillary" only if it has an "organic connection" to the venture's operations and serves to make the venture operate more efficiently or effectively. The effectiveness of a joint venture may be enhanced by restraints that reduce the incentive or opportunity for participants in the venture to expropriate or undermine its value.

<sup>50</sup> *Polk Bros. v. Forest City Enters.*, 776 F.2d 185 (7th Cir. 1985).

<sup>51</sup> *Id.* at 187.

<sup>52</sup> *Id.* at 190. The district court had found that the stores would not have been built without the restrictive covenant. *Id.*

<sup>53</sup> *Id.* at 189.

<sup>54</sup> *Dagher v. Saudi Refining Inc.*, 369 F.3d 1108, 1118 (9th Cir. 2004), *rev'd*, 547 U.S. 1 (2006).

<sup>55</sup> *Gen. Leaseways, Inc. v. Nat'l Truck Leasing Ass'n*, 744 F.2d 588, 595 (7th Cir. 1984) (Posner, J.) (holding that a territorial allocation was not ancillary to an arrangement by truck leasing companies for cooperative servicing each others' trucks). *See also Freeman v. San Diego Ass'n of Realtors*, 322 F.3d 1133, 1151 (9th Cir. 2003) ("The fixed support fee is not reasonably ancillary to the combined MLS database. Defendants offer no explanation for how it improves the efficiency of the MLS or has any effect at all beyond raising prices.")

<sup>56</sup> *PolyGram Holding, Inc. v. FTC*, 416 F.3d 29, 38 (D.C. Cir. 2005).

#### IV. WHAT IS THE “TEST” FOR ANCILLARITY?

After articulating the definition of an ancillary restraint quoted at the outset of this article,<sup>57</sup> Judge Bork added: “Of course, the restraint imposed must be related to the efficiency sought to be achieved. If it is so broad that part of the restraint suppresses competition without creating efficiency, the restraint is, to that extent, not ancillary.”<sup>58</sup> Similarly, Judge Taft indicated that a restraint is ancillary if “commensurate only with the reasonable protection of the covenant” and thus “reasonably necessary . . . to the enjoyment by the buyers of the property, good will, or interest in partnership bought.”<sup>59</sup> This “reasonably necessary” test for ancillarity finds much support in the modern case law.

In his dissent from the denial of certiorari in the *NASL* case, Justice Rehnquist asserted that: “The antitrust laws impose a standard of reasonableness, not a standard of absolute necessity.”<sup>60</sup> And the Eleventh Circuit recently held that: “Ancillary restraints are generally permitted if they are ‘reasonably necessary’ toward the contract’s objective of utility and efficiency.”<sup>61</sup> The “reasonably necessary” test also has been applied in Section 1 cases involving joint ventures but not expressly invoking the ancillary restraints doctrine.<sup>62</sup>

Also without expressly invoking the ancillary restraints doctrine, the *Antitrust Guidelines for Collaborations among Competitors (Guidelines)*, issued by the FTC and Department of Justice, state that the agencies apply the rule of reason to “agreements of a type that otherwise might be considered per se illegal, provided they are reasonably related to, and reasonably necessary to achieve procompetitive benefits from,” legitimate joint ventures.<sup>63</sup> The *Guidelines* explain that, in determining whether a restraint is “reasonably necessary,” the issue is “whether practical, significantly less restrictive means were reasonably available when the agreement was entered into.”<sup>64</sup> Critically, as Judge Easterbrook stressed in *Polk Bros.*, the need for the restraint is evaluated “at the time it was adopted.”<sup>65</sup>

Judge Bork provided one of the most thorough applications of the “reasonably necessary” test in *Rothery*, which concerned the activities of Atlas Van Lines, a nationwide interstate mover of household goods.<sup>66</sup> Throughout the country, Atlas employed independent agents for which it provided training, uniforms, and equipment used to supply the Atlas brand of interstate moving services. Unless barred by Atlas from doing so, many of its agents could have used what Atlas provided to supply intrastate moving services. To prevent such free riding, Atlas required its agents either to eschew any independent intrastate operations, or create new corporations to conduct intrastate business and not use any Atlas equipment, facilities, or services.<sup>67</sup> Judge Bork found that Atlas’ policy was “a classic attempt to counter the perceived menace that free riding poses,” and thus “reasonably necessary” to the joint venture.<sup>68</sup>

A restraint’s ancillarity is apt to be a disputed issue for trial. In the *Visa* case,<sup>69</sup> the two charge card joint ventures attempted to show that their rules prohibiting member banks from issuing American Express or Discover cards were reasonably necessary to promote “loyalty” and “cohesion.” After trial, the district court held that “defendants [had] not met their burden to come forward with a

57 See *supra* text accompanying note 6.

58 *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.*, 792 F.2d 210, 224 (D.C. Cir. 1986). See also *Schering-Plough Corp. v. FTC*, 402 F.3d 1056, 1073 (11th Cir. 2005) (“Naturally, the restraint imposed must relate to the ultimate objective, and cannot be so broad that some of the restraint extinguishes competition without creating efficiency.”), *cert. denied*, 126 S. Ct. 2929 (2006).

59 *United States v. Addyston Pipe & Steel Co.*, 85 F.2d 271, 281, 290-91 (6th Cir. 1898), *aff’d as modified*, 175 U.S. 211 (1899).

60 *NFL v. N. Am. Soccer League*, 459 U.S. 1074, 1079 (1982); see also *id.* at 1080 (arguing that the lower court erred by “ignor[ing] its own holding that the proper standard is that the constraint be ‘reasonably necessary’”).

61 *Schering-Plough*, 402 F.3d at 1072. See also *Freeman v. San Diego Ass’n of Realtors*, 322 F.3d 1133, 1151 (9th Cir. 2003) (stating that the issue is “reasonably ancillary to the legitimate cooperative aspects of the venture”); 11 HOVENKAMP, *supra* note 11, Paragraph 1912c2, at 332 (“one must consider whether the challenged restraint is necessary or makes and important contribution to” a legitimate arrangement).

62 See, e.g., *Law v. NCAA*, 134 F.3d 1010, 1019 (10th Cir. 1998) (inquiring whether the challenged conduct is “reasonably necessary to achieve the legitimate objectives”); *Orson, Inc. v. Miramax Film Corp.*, 79 F.3d 1358, 1367-68 (3d Cir. 1996) (inquiring whether the restraint is “reasonably necessary to achieve the stated objective”); *Thompson v. Metro. Multi-List, Inc.*, 934 F.2d 1566, 1581 (11th Cir. 1991) (inquiring whether the restraint is “reasonably necessary to the accomplishment of the legitimate goals and narrowly tailored to that end”).

63 FTC & U.S. DEP’T OF JUSTICE, ANTITRUST GUIDELINES FOR COLLABORATIONS AMONG COMPETITORS Section 1.2 (April 2000), *reprinted in* 4 Trade Reg. Rep. (CCH) Paragraph 13,160.

64 *Id.* Section 3.2.

65 *Polk Bros. v. Forest City Enters.*, 776 F.2d 185, 189 (7th Cir. 1985).

66 *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.*, 792 F.2d 210, 211 (D.C. Cir. 1986).

67 *Id.* at 211-13, 217, 221-23.

68 *Id.* at 223, 227.

69 *United States v. Visa U.S.A., Inc.*, 163 F. Supp. 2d 322 (S.D.N.Y. 2001), *aff’d*, 344 F.3d 229 (2d Cir. 2003), *cert. denied*, 543 U.S. 811 (2004).

valid procompetitive justification” for these rules, which had been shown to have anticompetitive effects.<sup>70</sup> The court found that the rules were not actually motivated by a desire to promote “cohesion” and that there was insufficient evidence that the rules, in fact, promoted “cohesion.”<sup>71</sup> Critically, the burden was placed on the defendants to demonstrate that their rules were “reasonably necessary,” and they failed to do so.

The rationale for allocating the burden in this manner is essentially that of *NCAA*, which held that a “naked restraint on price and output requires some competitive justification.”<sup>72</sup> Of course, ancillary restraints are, by definition, not naked, but restraints may be treated as naked until they are shown to be ancillary. Although never using the word “ancillary,” this was precisely Chief Judge Ginsburg’s approach in *PolyGram*:

If, based upon economic learning and the experience of the market, it is obvious that a restraint of trade likely impairs competition, then the restraint is presumed unlawful and, in order to avoid liability, the defendant must either identify some reason the restraint is unlikely to harm consumers or identify some competitive benefit that plausibly offsets the apparent or anticipated harm.<sup>73</sup>

Thus, a collateral restraint is ancillary if, and only if, it was reasonably necessary to make the venture operate more efficiently or effectively at the time the restraint was entered into. Defendants have the burden of demonstrating ancillarity, but that burden is not a heavy one. The more important of the two words in “reasonably necessary” is the first. Defendants are required to show only that the ability of their legitimate joint venture to function efficiently or effectively was compromised in a significant way that was sensibly addressed by the restraint.

## V. WHAT SHOULD ONE MAKE OF *TOPCO* AND *CITIZEN PUBLISHING*?

The Supreme Court’s 1972 decision in *Topco* applied the per se rule to a restraint collateral to a plainly procompetitive joint venture.<sup>74</sup> *Topco Associates* supplied its owners – about twenty-five small and medium-sized supermarket chains – with private label products (store brands). *Topco* entailed “no pooling of earnings, profits, capital, management, or advertising resources” but rather served only as a purchasing agent for many items the member supermarkets sold under brand names owned by *Topco*.<sup>75</sup> The Court explained that forming *Topco* allowed its members “to compete more effectively with larger national and regional chains.”<sup>76</sup> The Court also noted that “the profit” derived by the *Topco* members from its private label brands was “substantial” and selling those brands “improved the competitive potential of *Topco* members.”<sup>77</sup>

The distribution of the *Topco* brands was subject to complicated territorial licenses featuring some exclusive territories.<sup>78</sup> *Topco* argued that it “could not exist if territorial divisions were anything but exclusive,” so the restraint “actually increase[d] competition.”<sup>79</sup> Because it agreed, the district court found the restraints procompetitive and hence lawful.<sup>80</sup> The Supreme Court did not reject, or even consider, whether the territorial restraints were procompetitive, but instead held that they were per se violations of section 1 of the Sherman Act.<sup>81</sup>

Chief Justice Burger dissented,<sup>82</sup> viewing the “principal purpose” of the joint venture – “to make economically feasible” the sale of private label products – as “unquestionably lawful” and

70 *Id.* at 399.

71 *Id.* at 399-405.

72 *NCAA v. Bd. of Regents of the Univ. of Okla.*, 468 U.S. 85, 110 (1984).

73 *PolyGram Holding, Inc. v. FTC*, 416 F.3d 29, 36 (D.C. Cir. 2005).

74 *United States v. Topco Assocs.*, 405 U.S. 596 (1972). For an extraordinarily in-depth treatment of the case, see Peter Carstensen & Harry First, *Rambling Through Economic Theory: Topco’s Closer Look*, in *ANTITRUST STORIES* 171 (Eleanor M. Fox & Daniel A. Crane, eds. 2007).

75 *Topco*, 405 U.S. at 589-99.

76 *Id.* at 599; see also *id.* at 599 n.3 (detailing the advantages derived from selling private label products).

77 *Id.* at 600.

78 *Id.* at 601-03.

79 *Id.* at 605.

80 *Id.* at 605-06; *United States v. Topco Assocs.*, 319 F. Supp. 1031, 1042-43 (N.D. Ill. 1970).

81 *Topco*, 405 U.S. at 608-12. The Court maintained that it was well established that the per se rule applied to horizontal territorial restraints, *id.* at 608, but Chief Justice Burger in dissent made a strong argument that no prior decision had squarely held that a horizontal territorial restraint by itself constituted a per se violation, *id.* at 615-19.

82 Justice Blackmun concurred in the result, decrying the “anomalous” result of the decision which would be to “stultify” competition. *Id.* at 612-13. Justices Powell and Rehnquist did not take part.

characterizing the territorial restrictions as “minimal ancillary restraints that are fully reasonable in view of the principal purpose.”<sup>83</sup> He was inclined to agree with the district court that the restraints tended “positively to promote competition in the supermarket field and to produce lower costs for consumers,” and he saw no reason to adopt a rule that would “lead to the likely demise of” competition from the Topco products.<sup>84</sup>

The district court had found that “the cost of developing consumer acceptance for the Topco private brands was born by each member in its own territory.”<sup>85</sup> Consequently, the territorial restraints were a reasonable response to a free-rider problem. In his *Rothery* opinion, Judge Bork observed that Topco “was a contractual integration of legally independent businesses designed to achieve efficiencies unavailable to its members separately. . . . Topco had an ancillary horizontal restraint designed to make the integration more effective. . . . This restraint had a clear relationship to marketing effectiveness.”<sup>86</sup> In addition, the only competition potentially restrained was that in the sale of Topco private label products,<sup>87</sup> and those products would not have existed absent the joint venture. Per se treatment is not appropriate for a restraint on competition created by a joint venture,<sup>88</sup> so such a restraint should always be *treated* as ancillary, i.e., not analyzed separately from the formation of the joint venture.

Although the Supreme Court has not questioned the wisdom of *Topco*,<sup>89</sup> distinguished judges sitting on the courts of appeals have done so. Judge Bork’s *Rothery* opinion concluded that *Topco* had been “effectively overruled” by subsequent Supreme Court decisions involving joint ventures.<sup>90</sup> Judge Boudin opined that: “*Topco* may no longer be good law for its broader proposition that . . . a restraint is condemned *per se* even where it is ancillary to a productive joint venture.”<sup>91</sup> And Judge Posner declared that the *Topco* and like decisions “are dead as dodos.”<sup>92</sup>

The Supreme Court’s 1969 decision in *Citizen Publishing* also applied the per se rule to restraints associated with a joint venture.<sup>93</sup> In 1940 the only two daily newspapers in Tucson, Arizona established a joint venture company to produce and distribute both papers.<sup>94</sup> Although the two papers retained separate ownership as well as separate news and editorial operations, the joint operating agreement (JOA) between them eliminated economic competition. Subscription and advertising rates were set jointly, costs and revenues were pooled with profits distributed in a fixed ratio, and each paper agreed not to engage in any competing publishing activity in the area.<sup>95</sup>

More than twenty years after the formation of the joint venture, the government filed suit, alleging that specific provisions of the JOA were per se violations of Section 1. The district court granted the government’s motion for a preliminary injunction, enjoined the challenged provisions of the JOA, and required independent ownership of the two papers.<sup>96</sup> The government did not challenge the joint venture as a whole, however, and the district court found that the sharing of “circulation, advertising, and production departments has resulted in substantial cost savings.”<sup>97</sup> The injunction, therefore, permitted the papers to continue sharing these departments.

On direct appeal, the Supreme Court affirmed, holding that the Section 1 violations alleged by the Department and found by the district court were “plain beyond peradventure.”<sup>98</sup> The

83 *Id.* at 613-14.

84 *Id.* at 623-24.

85 *Topco*, 319 F. Supp. at 1042.

86 *Rothery Storage & Van Co. v. Atlas Van Lines, Inc.*, 792 F.2d 210, 225 (D.C. Cir. 1986). Similar views were expressed by BORK, *supra* note 5, at 274-79; RICHARD A. POSNER & FRANK H. EASTERBROOK, ANTITRUST 248-49 (2d ed. 1981).

87 Chief Justice Burger emphasized this point (see *Topco*, 405 U.S. at 613) but did not take the argument a step further.

88 The 1995 Intellectual Property Guidelines, *supra* note 19, make a similar point. In Section 4.1.2, they explain why a licensing arrangement does not present antitrust concerns by stating that it “does not diminish competition that would occur in its absence.”

89 The Court has never suggested that *Topco* is not good law and most recently cited the facts and holding of the case in *Palmer v. BRG of Georgia, Inc.*, 498 U.S. 46, 49 (1990).

90 *Rothery*, 792 F.2d at 229 (citing *Nw. Wholesale Stationers, Inc. v. Pac. Stationery & Printing Co.*, 472 U.S. 284 (1985); *NCAA v. Bd. of Regents of the Univ. of Okla.*, 468 U.S. 85 (1984); *Broad. Music, Inc. v. Columbia Broad. Sys., Inc.*, 441 U.S. 1 (1979)).

91 *Mass. Food Ass’n v. Mass. Alcoholic Beverages Control Comm’n*, 197 F.3d 560, 564 n.2 (1st Cir. 1999).

92 RICHARD A. POSNER, ANTITRUST LAW 189 n.62 (2d ed. 2001) (referring also to *United States v. Sealy, Inc.*, 388 U.S. 350 (1967)).

93 *Citizen Publ’g Co. v. United States*, 394 U.S. 131 (1969).

94 *Id.* at 133-34.

95 *Id.* at 134.

96 *United States v. Citizen Publ’g Co.*, 280 F. Supp. 978, 993-94 (D. Ariz. 1968).

97 *Id.* at 982.

98 *Id.* at 135.

only issue addressed by the Court at any length was whether the failing company doctrine could justify the arrangement. Sidestepping whether the district court was correct in refusing to consider such a defense because the joint venture was not a merger, the Court held instead that the newspapers could not carry the “burden of proving that the conditions of the failing company doctrine have been satisfied.”<sup>99</sup>

Congress prevented *Citizen Publishing* from affecting joint ventures in the newspaper industry by promptly creating an antitrust exemption for newspaper JOAs.<sup>100</sup> And the decision was only occasionally, and only briefly, mentioned by the courts until it was discussed extensively in *Dagher*.<sup>101</sup> Similarities between the two cases naturally led the plaintiffs in *Dagher* to rely heavily on *Citizen Publishing*, and the Supreme Court did not explain why their reliance was misplaced.

The Supreme Court’s *Dagher* decision mentions *Citizen Publishing* as case illustrating the proposition that the ancillary restraints doctrine governs “the validity of restrictions imposed by a legitimate business collaboration, such as a business association or joint venture, on nonventure activities.”<sup>102</sup> Since the Court found the restraints in *Citizen Publishing* per se illegal, the Court’s citation of the case in *Dagher* must imply both that condemned restraints were on non-venture activities and were non-ancillary.<sup>103</sup>

The Court’s decision in *Citizen Publishing* provides no indication that the Court considered the possibility that the restraints were ancillary, and petitioners did not invoke the ancillary restraints doctrine their brief.<sup>104</sup> That the restraints were non-ancillary, however, was implicitly argued in the government’s brief. The government explained that a JOA permitted “the realization of the benefits of operating economies without requiring the toleration of all anti-competitive effects” and that the asserted justification for the restraints was merely that the papers “needed the increased profits which the elimination of commercial competition between them would make possible.”<sup>105</sup> As noted above,<sup>106</sup> merely generating profits does not make a restraint ancillary.

It is more difficult to understand why the restraints in *Citizen Publishing* did not “involve the core activity of the joint venture itself.”<sup>107</sup> The Court’s decision in that case declared that the “purpose of the JOA was to end any business or commercial competition between the two papers.”<sup>108</sup> In this regard, the Court plainly was not viewing the restraints as separate from the JOA, but rather as core elements of it, and that certainly seems an apt characterization.

The proper analysis of *Citizen Publishing* within the framework of *Dagher* may be that restraints relating to “core activity of the joint venture” were sufficiently anticompetitive to render the joint venture as a whole unlawful under the rule of reason.<sup>109</sup> Declaring the joint venture as a whole unlawful, of course, would not imply that the papers could not form a somewhat different joint venture preserving the cost savings they had achieved, so the proper remedy would have been essentially that imposed by the district court and upheld on appeal. Although this analysis leads to the remedy actually applied, it does not involve the application of the per se rule, so this plainly was not

99 *Id.* at 136-39.

100 Newspaper Preservation Act, Pub. L. No. 91-353, 84 Stat. 466, codified at 15 U.S.C. Sections 1801-04.

101 *Dagher v. Saudi Refining Inc.*, 369 F.3d 1108, 1118-20, 1124 (9th Cir. 2004), *rev’d*, 547 U.S. 1 (2006).

102 *Texaco Inc. v. Dagher*, 547 U.S. 1, 7 (2006).

103 I previously argued that the Court’s analysis in *Citizen Publishing* treated the restraints as non-ancillary. See Werden, *supra* note 18, at 708 n.26.

104 The ancillary restraints doctrine was invoked by an amicus curiae brief. Brief of American Newspaper Publishers Assoc. as Amicus Curiae 54-58, *Citizen Publ’g Co. v. United States*, 394 U.S. 131 (1969).

105 Brief of the United States 36-37, *Citizen Publ’g Co. v. United States*, 394 U.S. 131 (1969). The district court had found that the cost-reducing joint printing and distribution of the two papers “does not depend upon the continuation” of the challenged restraints. *United States v. Citizen Publ’g Co.*, 280 F. Supp. 978, 992 (D. Ariz. 1968).

106 See *supra* text accompanying note 57.

107 *Dagher*, 547 U.S. at 7.

108 *Citizen Publ’g*, 394 U.S. at 134. In addition, the Department of Justice had argued that the profit pooling “was an integral part of the” JOA. See Brief of the United States 23, *supra* note 105.

109 Treating the JOA as a merger might be appropriate in view of the fact that it integrated the operations of the two papers and its purposes was to end commercial competition between them. Many transactions termed joint ventures by their participants have been treated as mergers under the antitrust laws. E.g., *United States v. Penn-Olin Chem. Co.*, 378 U.S. 158, 167-72 (1964); *FTC v. Warner Communications Inc.*, 742 F.2d 1156, 1159, 1163-65 (9th Cir. 1984); *United States v. Inaco, Inc.*, 704 F. Supp. 1409, 1414 (W.D. Mich. 1989). The federal enforcement agencies have stated that a joint venture “ordinarily” should be treated as a horizontal merger if “(a) the participants are competitors in that relevant market; (b) the formation of the collaboration involves an efficiency-enhancing integration of economic activity in the relevant market; (c) the integration eliminates all competition among the participants in the relevant market; and (d) the collaboration does not terminate within a sufficiently limited period by its own specific and express terms.” ANTITRUST GUIDELINES, *supra* note 63, at Section 1.3. Treatment as a merger would not affect the analysis or conclusions. See *United States v. Rockford Mem’l Corp.*, 898 F.2d 1278, 1282-83 (7th Cir. 1990).

the Court's analysis in *Citizen Publishing*.<sup>110</sup> Thus, it seems best to conclude that *Citizen Publishing* is no longer good law after *Dagher*.

## VI. CONCLUSION

The ancillary restraints doctrine holds that a restraint collateral to a legitimate joint venture, limiting the ability of participants to compete outside the venture, is assessed as part of a package that includes the formation of the venture, if at the time the restraint was adopted, the restraint had an organic connection to venture and was reasonably necessary to make the venture more efficient or effective in achieving its procompetitive purposes.

---

<sup>110</sup> The application of the per se rule to the joint venture as a whole would be proper only if it was viewed essentially a cartel, but that view was neither suggested by the Department of Justice nor adopted by the Supreme Court, and it is belied by the cost reductions found by the district court. See *supra* note 97 and accompanying text.



# GET-OUT-OF-JAIL-FREE CARDS: AMNESTY DEVELOPMENTS IN THE UNITED STATES AND CURRENT ISSUES

---

*Constance K. Robinson  
Kilpatrick Stockton  
Washington, DC*

## I. INTRODUCTION

First and foremost, antitrust amnesty programs are designed to help detect, deter and prosecute cartels by awarding “lenient” treatment to – not indicting – those who voluntarily report illegal activity.<sup>1</sup> As Assistant Attorney General Thomas O. Barnett has explained:

It is notoriously difficult to discover cartel behavior or, once discovered, to compile sufficient evidence to successfully prosecute cartel members in court. To penetrate the elaborate concealment strategies cartels use, prosecutors must have a tool to convert cartel members into cooperative witnesses, so that prosecutors can gain access to background information, testimony, and the documents that otherwise might be destroyed. Amnesty programs are such a tool.<sup>2</sup>

The United States’ current amnesty program has been very successful by any measure – the increased number of amnesty applications and cases, including internal investigations and prosecutions, and the higher amount of penalties obtained.<sup>3</sup> The number of applications now averages two per month, compared to one per year under the previous policy. Fines which used to be in the single digits have now been as high as \$500 million, and the Division has had fines imposed of over \$3 billion, over 90% of which were from international cartels.<sup>4</sup> And since imitation is the sincerest form of flattery, the fact that the rest of the world has followed the United States’ lead in adopting amnesty policies underscores the successfulness of the policy.<sup>5</sup>

In addition to their role in identifying and stopping cartels, leniency programs accommodate two fundamental, and sometimes conflicting, policy goals of criminal antitrust law: efficiency and equity. Where government investigations result in the expenditure of significant amounts of time and resources without yielding sustainable convictions or evidence sufficient to support such convictions, the criminal process is inefficient. Leniency policies remedy this inefficiency by encouraging participants in criminal antitrust violations to come forward to prosecutors as quickly as possible with high quality, direct information of the crime, which enables the government to reduce the time and increase the likelihood for obtaining a conviction.

Leniency also must balance the level of this encouragement against criminal law’s principles of fairness and retribution. Providing amnesty to a highly culpable ring-leader and coercer of a criminal antitrust conspiracy may provide prosecutors with an early bounty of highly useable evidence, but not without violating the criminal law’s sense of fairness by not appropriately punishing

<sup>1</sup> “Lenient” treatment is also called amnesty or immunity.

<sup>2</sup> “*Criminal Enforcement of Antitrust Laws: The U.S. Model*” by Thomas O. Barnett, President of the Fordham Competition Law Institute’s Annual Conference on International Antitrust Law and Policy, New York, New York, (September 14, 2006) (<http://www.usdoj.gov/atr/public/speeches/218336.htm>.)

<sup>3</sup> See R. Hewitt Pate, International Anti-Cartel Enforcement, address before ICN Cartels Workshop (Nov. 21, 2004) at 4 (<http://www.usdoj.gov/atr/public/speeches/206428.htm>).

<sup>4</sup> See Scott D. Hammond, *A Review of Recent Cases and Developments in the Antitrust Division’s Criminal Enforcement Program* 14 & n.9 (Apr. 24, 2002) (<http://www.usdoj.gov/atr/public/speeches/10862.htm>).

<sup>5</sup> “*An Overview of Recent Developments in the Antitrust Division’s Criminal Enforcement Program*,” by Scott D. Hammond, Deputy Assistant Attorney General, Before the American Bar Association (January 10, 2005) (<http://www.usdoj.gov/atr/public/speeches/213247.htm>.)

the criminals most responsible for the violation. Alternatively, leniency policies that weigh the balance too heavily in favor of fairness run the risk of reducing the incentives for confessors to come forward. Finding the balance between efficiency and fairness has marked the evolution of the U.S. Department of Justice's leniency policies.

This paper begins by reviewing the history of that evolution in the United States' leniency policy, from its beginning in 1978 to the current corporate and individual leniency policies, as well as the "Amnesty Plus" and "Penalty Plus" programs, and the advantages of cooperation even where a confessor is not the first to approach the agencies. Next, the paper reviews some of the more relevant recent developments affecting the application and scope of leniency in the United States, including: the *Stolt-Nielsen* investigation, the U.S. Justice Department policy referred to as the McNulty Memorandum, the Antitrust Criminal Penalty Enhancement and Reform Act, the latest amendments to the Sentencing Guidelines and the recent debate before the Antitrust Modernization Commission. A companion paper by Tim Fraser assesses the emergence, and convergence, of leniency policies in other jurisdictions around the world. Although this paper is limited to amnesty under the antitrust laws, criminal activity may implicate other types of laws and other types of amnesty programs.<sup>6</sup>

## II. HISTORY OF THE ANTITRUST DIVISION'S LENIENCY POLICY

The Antitrust Division unveiled the world's first antitrust leniency policy in 1978. The policy underwent three substantial revisions in 1993 and was supplemented with a leniency policy for individuals in 1994. Recently, the Division enhanced the leniency policy's leverage with the deployment of the amnesty plus and penalty plus programs. Despite the revisions and enhancements, the basic 1978 design for the program remains intact today.

### Division's 1978 Amnesty Policy

On October 4, 1978, the Assistant Attorney General for the Antitrust Division of the Department of Justice, John Shenefield, announced the Division's prosecutorial amnesty policy.<sup>7</sup> Under the policy, corporations or officers who voluntarily reported their criminal violations of the antitrust laws to the Division *prior* to its detecting the conduct were *eligible* for discretionary leniency from prosecution. The policy afforded leniency only to the first "whistleblowing conspirator" to come forward, whereas any subsequent confessors were prosecuted but given more lenient sentencing recommendations. For a confession to meet the policy's standards, it had to be a truly corporate act, not simply a confession by an individual executive or officer.

Upon receiving a pre-investigation confession, the Division applied five additional factors in deciding whether to grant the confessor leniency: (1) whether the Division could have reasonably expected that it would have become aware of the conspiracy in the near future, absent the corporation's confession; (2) whether the corporation, on discovering the illegal activity took quick, effective action to end its participation in the conspiracy; (3) whether the corporation's confession was made with candor and completeness and whether it continued to cooperate with the Division throughout the investigation; (4) the nature of the violation and the confessing party's role in it, *e.g.*, whether the corporation coerced co-conspirators into participating in the conspiracy, whether its conduct had an actual exclusionary effect on others in the marketplace, whether it originated the scheme, and whether the scheme harmed competition; and (5) whether the corporation had made or stated its intent to make restitution to injured parties.

The 1978 policy did not elaborate on how the Division would weigh the factors, the first three of which enhance the efficiency of the leniency program by increasing the incentives for confessors to act quickly and effectively, and the last two of which are geared towards maintaining equitable principles in granting leniency by not letting the most culpable or least contrite off the

<sup>6</sup> For example, there are many parallels between the United States' Corporate Leniency Policy and the Security and Exchange Commission's Seaboard Report, which outlines the Commission's enforcement policy with respect to voluntary disclosures and cooperation. *See* Report of Investigation Pursuant to Section 21(a) of the Securities and Exchange Act of 1934 and Commission Statement on the Relationship and Cooperation to Agency Enforcement Decisions, Exchange Act Release No. 44, 969, Accounting and Auditing Enforcement Release No. 1,470 (Oct. 23, 2001); *see also* Statement of the Securities and Exchange Commission Concerning Financial Penalties (2006).

<sup>7</sup> *See* 4 Trade Reg. Rep. (CCH) Paragraph 13,112 (2001).

hook. This lack of transparency and certainty resulted in the policy's failure to be as successful as anticipated, producing only one confessor per year and failing to uncover a single international cartel.<sup>8</sup>

### 1993 – Corporate Leniency Policy

On August 10, 1993, the Division announced three substantial revisions to its Corporate Leniency Policy designed to provide transparency and certainty and to increase incentives to cooperate. First, where leniency under the old policy had been discretionarily given to corporations confessing their culpability prior to the Division detecting the violation, it is now *assured* under the new policy. Second, a corporate confessor reporting a violation during an ongoing investigation may be eligible for leniency under the current policy, where it would not have been under the 1978 policy. Finally, officers, directors and employees coming forward with a corporation qualifying for automatic amnesty are also assured amnesty, provided they cooperate with the investigation. There had been no such guarantee under the previous leniency policy.

To qualify for “Part A,” or automatic leniency, a corporate confessor reporting illegal activity must do so before an investigation has been opened and must also meet six additional conditions. First, the Antitrust Division must not have received information about the violation before the corporate confessor comes in. Second, the confessor must have promptly and effectively terminated its participation in the illegal activity upon discovering it. Third, the report of the violation must be full, complete and candid, and the corporation must continue to cooperate with the Division's investigation. Fourth, the confession must be a true corporate act, not simply confessions of individuals. Fifth, the corporation must make restitution to the injured parties where possible. Finally, the confessor must not have been the coercer, ring-leader or originator of the illegal activity. The Division has since clarified this final condition to disqualify confessors only if they are the lone ringleaders or organizer of the conspiracy.<sup>9</sup> While these conditions are essentially similar to those evaluated in the 1978 policy, the current policy clarifies that all of them must be met to qualify.

If a corporation qualifies for automatic or “Part A” leniency, so do its officers, directors and employees, provided that they admit their involvement in the illegal activities with candor and completeness and continue to cooperate with the Division in its investigation. If a corporation does not qualify for automatic immunity, it may still be eligible for discretionary, or “Part B” immunity.

Under “Part B” immunity, a corporation's confession that comes after the Division has received information about the violation from another source, whether or not an investigation has begun, or under circumstances where the corporation was a central figure in the conspiracy may qualify if it meets all the other conditions for automatic leniency, and if the following additional conditions are satisfied: 1) it is the first confessor to come forward and qualify for leniency; 2) at the time of the confession, the Division has not collected evidence sufficient to produce a sustainable conviction;<sup>10</sup> and 3) in the exercise of its discretion, the Division determines that it would not be unfair to grant the corporate defendant leniency given its role in the illegal activity, the nature of the violation, and when the confessor came forward.

In exercising its discretion to grant leniency, the Division considers how long it took the corporation to come forward and the centrality of the corporation's role in the conspiracy. The Division is much more likely to find it fair to grant a corporate defendant leniency if the corporation comes in before an investigation has begun. The stronger the case the Division has developed independent of the confession, the less likely it is that the Division will find it fair to grant the confessor leniency.

If a corporation qualifies for discretionary leniency, its employees, officers and directors who assist the Division with its investigation are also eligible for leniency. The Division will consider

<sup>8</sup> See J. Anthony Chavez, *The Carrot and Stick Approach to Antitrust Enforcement*, 1542 Prac. L. Inst. / Corp. L. & Prac. Course Handbook Series 519, 551 (PLI Order No. 8736) (2006); Scott D. Hammond, Department of Justice, *Cornerstones of An Effective Leniency Program* (presented before the ICN Workshop on Leniency Programs, Sydney, Australia) (Nov. 22-23, 2004) (<http://www.usdoj.gov/atr/public/speeches/206611.htm>).

<sup>9</sup> *Id.* fn. 4.

<sup>10</sup> Congress' recent authorization of wiretapping in criminal antitrust investigations, see 18 U.S.C. Section 2516(1)(r), may make it more likely in future cases that the Division will have evidence sufficient to sustain a conviction without the corporate confessor.

granting these people leniency on the same basis as if they had approached the Division as individuals seeking leniency.

### **1994 – Leniency Policy for Individuals**

Exactly one year after announcing its Corporate Leniency Policy, the Antitrust Division announced its Leniency Policy for Individuals on August 10, 1994. Under the policy, individuals may obtain leniency from the Division without first waiting for their corporation to seek it. The Division will grant automatic leniency to individuals who report a criminal antitrust violation before the Division has begun an investigation and before it has received information about the violation. Additionally, to qualify for automatic leniency, the individual must confess with candor and completeness and cooperate with the Division's investigation. A final condition to obtaining automatic leniency is that the individual confessor was not the coercer, leader or originator of the criminal activity. Individuals who do not qualify for automatic leniency under the Leniency Policy for Individuals may still be considered by the Division in its discretion for either statutory or informal immunity.

### **Amnesty “Plus”/Penalty “Plus”**

In addition to its formal corporate and individual leniency policies, the Antitrust Division has developed two important and supplemental strategies for getting confessors to disclose information regarding antitrust conspiracies. The first strategy has been dubbed “Amnesty Plus.” Under this policy, an applicant who may not qualify for amnesty from prosecution for one criminal offense may still be eligible for amnesty from prosecution by reporting a second criminal offense “plus” a reduction in the criminal penalty for the first offense. As with the revisions to the Corporate Leniency Policy, the Amnesty Plus program appears to have been successful. As of November 2005, the Division had “roughly 56 sitting grand juries investigating suspected international cartel activity. Nearly half of th[o]se investigations were initiated by evidence obtained as a result of an investigation of a completely separate industry.”<sup>11</sup>

Where Amnesty Plus provides incentives to inform the Division of a criminal violation, the “Penalty Plus” policy provides disincentives for not coming forward. A corporation that fails to provide information under the Amnesty Plus program of the second violation runs the risk of facing substantially higher fines should the second offense be discovered by the Division. In this circumstance, it is the Division's policy to seek an upward departure from the Sentencing Guidelines, which could mean a potential fine as high as 80 percent or more of the volume of affected commerce.

### **Second-In**

Second-in confessors, which are the second companies to confess their involvement in an illegal antitrust conspiracy after a first qualifying company comes forward under the Division's leniency program, are the primary beneficiaries of the Amnesty Plus policy. A second-in confessor who is ineligible for Amnesty Plus, because the Division has already opened an investigation in the second market, may be offered amnesty on the condition that it covertly assists the Division in its investigation of the second market.

In addition to increasing the probability of qualifying for Amnesty Plus or affirmative amnesty, there are four other significant benefits to being a second-in company. First, second-in companies may have their criminal penalty reduced by limiting the scope of the affected commerce used to calculate the penalty. The reason is that it is the Division's practice not to include self-incriminating information provided by second-in companies in calculating the applicable Sentencing Guidelines range for the company's volume of affected commerce. Second, second-in companies may obtain substantial cooperation discounts applied to the bottom of the Guidelines range on the order of 30% to 35%. Confessors following second-in companies may still be eligible for cooperation

<sup>11</sup> Scott D. Hammond, U.S. Dep't of Justice, *An Update of the Division's Criminal Enforcement Program* (Nov. 16, 2005) (<http://www.usdoj.gov/atr/public/speeches/213247.htm>).

discounts, albeit in smaller amounts. A third, and related, benefit is that second-in companies can assure a low-starting point in which to apply the cooperation discount. A second-in company is generally eligible for a starting point at the minimum Guidelines range, whereas subsequent confessors generally start well above this point. The fourth benefit of being a second-in company is securing more favorable treatment for its executives and employees, both by minimizing the number of them subject to prosecution and by maximizing the favorability of individual plea agreements.

### III. RECENT DEVELOPMENTS AFFECTING THE DIVISION'S LENIENCY POLICY

There have been a number of recent judicial and regulatory developments that affect the interpretation and strength of the Division's leniency policies. Recent judicial developments have both strengthened and weakened the Government's flexibility in handling leniency agreements.

*Stolt-Nielsen*, which has been much talked about but may not have a significant effect on the actual application of the leniency programs because it may be limited to the unusual facts of the case, highlights the risks confessors face when the Government has the power to withdraw its conditional leniency. Judicial retrenchment of the Justice Department's Thompson Memorandum, on the other hand, may substantially limit the type of "cooperation" the Government may demand from corporate confessors seeking leniency.

Recent regulatory and legislative developments, such as the Antitrust Criminal Penalty Enhancement and Reform Act and amendments to the Sentencing Guidelines, have increased the incentives to prospective criminal defendants of seeking leniency. And the work of the Antitrust Modernization Commission may result in future legislative changes.

#### **Applying the Division's Corporate Leniency Policy – *Stolt-Nielsen***

In March 2004, *Stolt-Nielsen* became the first company to have its conditional leniency revoked by the Division since the introduction of the 1993 Corporate Leniency Policy.<sup>12</sup> Although *Stolt-Nielsen S.A.* turns on a unique fact-pattern, it is also an example of how stringently the Division will apply the requirement that corporations promptly and effectively cease their participation in the illegal activity being reported.

In November 2002, an antitrust lawyer, John Nannes, met with *Stolt-Nielsen's* Chairman, Samuel Cooperman regarding a February 2002 memorandum written by *Stolt's* former in-house counsel, Paul O'Brien, which outlined potential evidence of a criminal antitrust violation.<sup>13</sup> Nannes was then retained by *Stolt* and authorized to begin the process of obtaining leniency for the company from the Antitrust Division. He began the process by asking his former Division colleague, James Griffin, then Deputy Assistant Attorney General for the Antitrust Division, whether the Division had opened up an investigation into the parcel tanker industry. Although not using any names explicitly, Griffin intimated that he had heard that O'Brien had been fired by *Stolt* for authoring the memorandum regarding potential antitrust violations, suggesting both that *Stolt* would be ineligible for amnesty under the program because the Division had already learned of the violations and that if he were fired for disclosing violations the company may not have promptly addressed the violations. (Apparently, the *Wall Street Journal's* publishing of an article on November 22, 2002, that described O'Brien's wrongful termination suit against *Stolt*, caused the Division to open its investigation of the parcel tanker industry.)<sup>14</sup>

In December 2002, Nannes met with the Division and informed it that even though O'Brien had been fired, *Stolt* followed O'Brien's recommendations for implementing an antitrust compliance program. Griffin advised Nannes that if *Stolt* had not in fact withdrawn from the conspiracy, it would not be eligible for amnesty under the program. Nannes stated that *Stolt* had taken steps to withdraw its participation, but he did not go so far as to say that *Stolt's* participation

12 U.S. Dep't of Justice, Press Release, *Stolt-Nielsen S.A. Indicted on Customer Allocation, Price Fixing, and Bid Rigging Charges for Its Role in An International Parcel Tanker Shipping Cartel* (Sept. 6, 2006).

13 *Stolt-Nielsen S.A. v. United States*, 352 F. Supp. 2d 553, 556 (E.D. Pa. 2005).

14 See *Stolt-Nielsen S.A.*, 352 F. Supp. 2d at 565.

ended in March 2002, which is the month after O'Brien's memorandum. (In fact, it appears that Stolt continued to participate in the conspiracy until November 2002, nine months after O'Brien's memorandum.) Shortly after the meeting, Nannes received a marker from the Division, which reserved Stolt's place as first in line for leniency eligibility, and began his internal investigation of Stolt and the allegations in O'Brien's memorandum.

In January 2003, Nannes made a proffer to DOJ, which included the fact that Stolt had participated in a market allocation scheme with its competitors. The date on which Stolt had withdrawn from the conspiracy was not apparently discussed. Soon after the proffer, Griffin sent Nannes a corporate leniency letter, which was later modified to cover all of Stolt's subsidiaries and affiliates. The letter represented that Stolt "took prompt and effective action to terminate its part in the anticompetitive activity being reported upon discovery of that activity," but did not mention the month in which it believed Stolt's participation ended. Stolt then provided documents and other information to the government, enabling it to obtain convictions for the other industry participants. Three months after the leniency letter was issued, however, the Division temporarily suspended Stolt's amnesty because it believed Stolt continued to participate in the market allocation scheme in the second half of 2002. In February of the following year, Stolt filed suit seeking an injunction to force the Division to honor the January 2003 letter agreement not to prosecute Stolt. The month after the suit, the Division formally withdrew Stolt's amnesty and notified Stolt that it would seek an indictment, which the Division later agreed to stay until Stolt's injunction suit was resolved. The Division's belief was that O'Brien's February 2002 memorandum constituted Stolt's discovery of the activity and that waiting until November to cease participation did not constitute prompt and effective termination. The Division believed that even though its letter did not state a date certain, the principle underlying the concept of timeliness was clear from its published policy.

The District Court ruled in favor of Stolt on the grounds that the Division's understanding of when Stolt withdrew from the conspiracy was not part of the amnesty agreement as it was not in the four corners of the contract, which had an integration clause expressly excluding outside terms from the agreement.<sup>15</sup>

The District Court's opinion was overruled by the Third Circuit in March 2006<sup>16</sup> holding that the lower court was without jurisdiction to enjoin the Government from seeking an indictment against Stolt.<sup>17</sup> The lower court lacked authority to enforce the agreement because immunity agreements protect defendants against conviction, not indictment, and thus did not implicate Stolt's constitutional rights, the only circumstance where the judiciary might enjoin prosecutions.<sup>18</sup>

A petition for a writ of certiorari was filed on July 20, 2006, along with an application to Justice Souter to recall and stay the mandate of the Court of Appeals pending the Court's disposition of the certiorari petition. Justice Souter denied the application. Stolt re-filed the application with Justice Stevens who referred it to the full Court. On August 21, 2006, the Court denied petitioners' application. Stolt, along with two subsidiaries and two executives, was then indicted by a federal grand jury.<sup>19</sup> Stolt filed a motion to dismiss the indictment on the basis of the leniency agreement on November 22, and evidentiary hearings have been held since May 30, 2007. The district court has not yet ruled on the motion to dismiss.

While it may appear that after this opinion there is nothing, other than trust, to guarantee confessors that the government will adhere to the letter of agreements reached under the Leniency Policy, there may be further developments as the trial court considers the effect of the letter on the possibility of conviction. Additionally, it is important to recognize that the Division has a strong incentive to honor its leniency agreements because the effectiveness of its programs depends on a high level of certainty. Thus, it is unlikely that *Stolt-Nielsen* will have a significant impact on the operation of the leniency programs. But the case highlights the risks associated with conditional leniency and it may cause some nervousness on the part of those in or seeking to enter the amnesty process.

<sup>15</sup> *Id.* at 562.

<sup>16</sup> *See Stolt-Nielsen, S.A. v. United States*, 442 F.3d 177, 187 (3d Cir. 2006), cert. denied, 127 S.Ct. 494 (2006).

<sup>17</sup> *Id.* at 187.

<sup>18</sup> *Id.* at 184-86.

<sup>19</sup> *See* U.S. Dep't of Justice, *supra* note 12.

## The Thompson Memo

On January 20, 2003, Larry Thompson, Deputy Attorney General of the United States, authored a memorandum binding on all federal prosecutors called *Principles of Federal Prosecution of Business Organizations* (hereinafter, the “Thompson Memo”), which was essentially identical to a memorandum authored by Eric Holder, Deputy Attorney General in 1999. The Thompson Memo required that federal prosecutors consider the advancing of legal fees by corporations to their employees under investigation by the Department of Justice as a factor weighing in favor of indicting the corporate defendant. Historically, most corporations under investigation for criminal wrongdoing extend attorneys’ fees to their principals and employees who are also being investigated in their individual capacities. After the Thompson Memo, corporations seeking lenient treatment from Justice Department prosecutors run the risk that, if they advance legal fees to their employees who are also under investigation, they may not qualify for leniency. Commentators and bar associations have roundly criticized this aspect of the Thompson Memo as an infringement on individual defendants’ Sixth Amendment right to counsel – by depriving individual defendants of the means to pay for their own defense – and due process rights under the Fifth Amendment – by essentially voiding defendants’ contractual rights to attorney fees. The criticism gained judicial support when the Thompson Memo was found to infringe on individual defendants’ right to a fair trial and to effective assistance of counsel by Judge Lewis Kaplan in *United States v. Stein*, 435 F.Supp.2d 330, 367 (S.D.N.Y. 2006) (*Stein I*). In a second ruling regarding the same KPMG tax fraud prosecution (*Stein II*), Judge Kaplan suppressed certain KPMG employee defendants’ proffered statements to the DOJ, finding that “the Thompson Memorandum and the actions of the USAO, quite deliberately coerced . . . KPMG to pressure its employees to surrender their Fifth Amendment rights” by inducing KPMG to threaten to terminate payment of attorneys’ fees for non-cooperative employees.<sup>20</sup>

## The McNulty Memo

On December 12, 2006, then U.S. Deputy Attorney General Paul J. McNulty announced a new DOJ corporate prosecutorial policy (hereinafter, the “McNulty Memo”) that would replace the Thompson Memo. The change was spurred by criticism from the business community and the defense bar of both the frequency and ease with which DOJ prosecutors demanded privileged material under the Thompson policy, the constitutional issues identified in *Stein I* and *Stein II* with respect to attorneys’ fees, and by proposed legislation by Senate Judiciary Committee Chairman Arlen Specter (R-Pa), The Attorney Client Privilege Protection Act.<sup>21</sup> The McNulty Memo changed the Thompson Memo in two significant ways.

First, the new policy places more stringent limitations on the procurement of privileged material. Prosecutors must demonstrate a legitimate need for privileged information, and any request stemming from the need is subject to an extensive approval process. The approval process depends on the type of waiver requested. Specifically, a request for waiver of claims of attorney-client privilege or attorney work product involving *purely factual* information (“Category I”) must be authorized in writing by the United States Attorney, or (for an antitrust violation) by the Assistant Attorney General for the Antitrust Division.<sup>22</sup> Only if Category I material gives an incomplete basis to conduct a thorough investigation should DOJ request a waiver of attorney-client communications or *non-factual* attorney work product (“Category II”). Waiver for Category II material should only be requested in rare circumstances and requires written approval from the United States Attorney or (for an antitrust violation) the Assistant Attorney General for the Antitrust Division, and the Deputy Attorney General.<sup>23</sup> Furthermore, and more important for corporate defendants, if the corporation decides not to waive its privilege to Category II information upon request from the government, the revised policy instructs prosecutors to exclude the denial from their charging decision-making process.<sup>24</sup> The

20 *United States v. Stein*, 440 F.Supp.2d 315, 337 (S.D.N.Y. 2006). Rather than dismiss the government’s indictment, Judge Kaplan opened a parallel civil proceeding to resolve the KPMG employees’ contract claims against KPMG. *Stein I*, 435 F. Supp.2d at 377-378. The Second Circuit then rejected the novel remedial procedure, vacating the order asserting ancillary jurisdiction as beyond the jurisdiction of the district court. *Stein v. KPMG, LLP*, 2007 WL 1487822 (2d Cir., May 23, 2007). The Second Circuit did not consider the constitutional issues decided by the district court.

21 Carrie Johnson, *Higher Hurdles Set in Corporate Crime Cases*, The Washington Post, Dec. 13, 2006 at D1.

22 Memorandum from Paul J. McNulty, *Principles of Federal Prosecution of Business Organizations*, to the U.S. Dept’t of Justice (Dec. 12, 2006) at 9, 11.

23 *Id.* at 10.

24 *Id.*

DOJ can, however, favorably consider a corporation's acquiescence to a waiver request in determining whether it has cooperated with a government investigation.

Second, the new policy instructs prosecutors not to consider a corporation's payment of its employees' attorneys' fees in deciding whether to pursue a conviction, unless the totality of the circumstances shows that payment of the fees was meant to hinder government investigation. If this unusual finding is made, corporate payment of attorney's fees may be considered in the charging decision only if authorized by the Deputy Attorney General.

### **The Antitrust Criminal Penalty Enhancement and Reform Act of 2004**

Increasing the incentives for participating in the Division's leniency programs, the Antitrust Criminal Penalty and Reform Act of 2004 imposes significantly higher criminal antitrust penalties for both businesses and individuals. The Act raised the maximum statutory criminal fine for corporations from \$10 million to \$100 million and for individuals from \$350,000 to \$1 million. (The alternative fine provision, 18 U.S.C. Section 3571(d), under which the Division can seek twice the gain or twice the loss remains unchanged; it is through use of this provision that many of the record settlements have been obtained.) Prison sentences under the Act were also increased, from a maximum prison sentence of 3 years to 10 years.

In addition to raising the stakes for convictions, the Act limits a cooperating defendant's exposure from private litigation to single, not the standard treble, damages, an extremely significant incentive to seek leniency. (The detrebling feature will sunset in 2009.) The Act defines cooperating with the government as: (1) providing a full account of all facts relevant to the civil action; (2) furnishing all documents relevant to the civil action; and (3) making oneself available for interviews, depositions and testimonies in connection with the civil action.

### **Compliance Programs and Sentencing**

Effective November 1, 2004, the United States Sentencing Commission amended the Sentencing Guidelines to provide that corporations employing an "effective" compliance program are eligible for a downward departure in sentencing.<sup>25</sup> The amendments were promulgated in response to Section 805(a)(2)(5) of the Sarbanes-Oxley Act of 2002, Public Law 107-204, which required the Commission to ensure that the Guidelines "are sufficient to deter and punish organizational criminal misconduct." These amendments have led to more pervasive and robust corporate compliance programs across the country.

Under Section 8B2.1 of the Guidelines, for an organization to have an effective compliance program, it shall: "(1) exercise due diligence to prevent and detect criminal conduct; and (2) otherwise promote an organizational culture that encourages ethical conduct and a commitment to compliance with the law." The program does not necessarily need to either detect or prevent criminal activity to be "effective," but the Guidelines identify seven minimum requirements for a compliance program to be deemed effective.

First, an effective program must have established "standards and procedures to prevent and detect criminal conduct."<sup>26</sup> This requirement shares parallels with the Department of Justice's leniency policies. Second, the control group or governing authority of the corporation must know about how the compliance program works and must also oversee the implementation and effectiveness of the program.<sup>27</sup> The oversight and responsibility for the compliance program shall be assigned to a "high-level" individual within the corporation,<sup>28</sup> whereas the individual with operational responsibility for the program must report periodically to the "high-level" personnel and the governing authority of the corporation regarding the program's effectiveness.

<sup>25</sup> United States Sentencing Commission, Guidelines Manual Section 8B2.1 (Nov. 1, 2004).

<sup>26</sup> *Id.* at Section 8B2.1(b)(1).

<sup>27</sup> *Id.* at Section 8B2.1(b)(2)(A).

<sup>28</sup> *Id.* at Section 8B2.1(b)(2)(B).



Third, the corporation shall not knowingly include in its control group any criminals or individuals whose conduct is inconsistent with ethics or an effective compliance program. Fourth, the corporation must train its governing authority, its high-level personnel, its substantial authority personnel, its employees and its agents with respect to its compliance standards and procedures. Fifth, the corporation must take reasonable steps to: (1) ensure that its program is being followed, “including monitoring and auditing to detect criminal conduct;” (2) evaluate its program’s effectiveness; and (3) publicize a system by which employees may report violations and seek guidance regarding the program.

Sixth, the corporation’s program must be promoted throughout the program by both providing incentives to comply as well as discipline for criminal conduct or failing to take reasonable steps to avoid criminal conduct. Finally, if the corporation detects criminal conduct, it must respond quickly and take measures to correct its compliance program and procedures.

In implementing the seven requirements outlined by the Guidelines, a corporation must periodically assess the risk that criminal conduct will occur, including assessing both the magnitude of the potential criminal conduct and its likelihood of occurring given the company’s particular business. “For example, an organization that, due to the nature of its business, employs sales personnel who have flexibility to set prices shall establish standards and procedures designed to prevent and detect price-fixing.”<sup>29</sup>

In deciding how exactly to meet these requirements, corporations are to consider: “(i) applicable industry practice or the standards called for by any applicable governmental regulation; (ii) the size of the organization; and (iii) similar misconduct.” Should the corporation fail to adhere to applicable industry standards of practice or relevant governmental regulation, that weighs against finding that the corporation had in place an effective compliance program.

The larger corporation will more likely have to make significant investments in a more formal compliance program for its program to be deemed “effective.” The smaller corporation, on the other hand, may satisfy the Guideline’s requirements by having its governing authority managing the company’s compliance efforts, by using informal staff meetings to train employees regarding compliance, by using its existent staff rather than employing specialized compliance staff, and by modeling its compliance efforts on best-practices developed by other organizations. The Guidelines presume that a company does not have an effective compliance program if criminal conduct recurs.

### **Other Potential Developments on Sentencing**

On April 2, 2007, The Antitrust Modernization Commission (“AMC”) issued its Report and Recommendations (“Report”) regarding statutes potentially in need of Congressional revision, including criminal penalty statutes. The first recommendation concerns whether 18 U.S.C. Section 3571(d), the alternative fine provision, should remain applicable to Sherman Act prosecutions. That subsection provides that: “[i]f any person derives pecuniary gain from the offense, or if the offense results in pecuniary loss to a person other than the defendant, the defendant may be fined not more than the greater of twice the gross gain or twice the gross loss, unless imposition of a fine under this subsection would unduly complicate or prolong the sentencing process.” Under this provision, corporate antitrust fines may exceed the statutory maximum of \$100 million, another reason why corporations have an incentive to seek amnesty. The AMC Report opposed amending Section 3751(d) to make it inapplicable to Sherman Act prosecutions, finding “nothing unique about antitrust offenses that justifies their being carved out or otherwise exempted” from the provision.<sup>30</sup>

The AMC also debated whether the ambiguity in language under Section 3751(d) about whether a fine should be calculated based on the individual defendant’s sales or whether it should be based on the sales of the cartel as a whole should be clarified. The commissioners ultimately voted not

<sup>29</sup> *Id.* at Section 8B2.1 cmt. 6(A)(ii).

<sup>30</sup> Antitrust Modernization Commission, Report and Recommendations 298 (2007).

to change the basis for calculating fines for Sherman Act violations, recommending that such “interpretive questions regarding the statute be left to courts to resolve in the context of actual cases.”<sup>31</sup>

The final issue debated by the AMC was the Sentencing Guidelines’ use of an estimate of the harm caused by an antitrust violation set at 20% of the amount of commerce affected by the violation. The specific question posed by the AMC was “whether the existing proxy is empirically sound and accurately reflects the best estimate of typical harm in antitrust cases.”<sup>32</sup> The commissioners recommended that Congress “encourage the Sentencing Commission to reevaluate and explain” the presumptive 20% harm proxy in light of modern economic understanding and analysis.<sup>33</sup> The commission also recommended that the Sentencing Commission amend the Guidelines to make explicit that the 20% harm proxy (or any revised proxy) may be rebutted by proof of a preponderance of the evidence that the actual amount of overcharge was higher or lower, where the difference would materially change the base fine.<sup>34</sup>

While it is far too early to predict the effect of the AMC’s report, to the extent fine levels increase through judicial interpretations of the alternative fine provision of Section 3751(d), the incentive for a corporation to seek amnesty should increase.

---

31 *Id.* at 299.

32 *Id.* at 301.

33 *Id.* at 300.

34 *Id.* at 295.

# IF IT TAKES TWO TO TANGO, DO THEY CONSPIRE?: TWOMBLY AND STANDARDS OF PLEADING CONSPIRACY

---

*Richard A. Duncan and Brian S. McCormac\**  
*Faegre & Benson LLP*  
*Minneapolis, MN*

The potentially malign results to competition of parallel decisions made by members of an oligopoly have long drawn academic attention. The two poles of the debate have conventionally been defined by Professor Donald Turner on the one hand, and then Professor Richard Posner on the other. Professor Turner argued that Section 1 of the Sherman Act, 15 U.S.C. Section 1, should generally not provide a remedy for such actions absent additional evidence of an agreement among the parties to refrain from competing. Professor (now Judge) Posner argued that Section 1 in principle covered such conduct, observing that in an oligopolistic setting the economic outcomes of “conscious parallelism” can mimic those that result from explicit cartels, and should receive similar legal treatment.<sup>1</sup>

The risks of anticompetitive behavior by firms in highly concentrated markets, or oligopolies, have been problematic for antitrust law and policy.<sup>2</sup> In an oligopoly, there are few competing firms, and economists have argued that these firms can increase prices above competitive levels without expressly communicating with each other.<sup>3</sup> Because the principal United States antitrust law dealing with collusive conduct among multiple firms, Section 1 of the Sherman Act, requires proof of agreement, or “conspiracy,” antitrust law has proven ill-suited to address the anticompetitive impact of parallel conduct by firms in oligopolies.<sup>4</sup>

In a series of cases decided between 1939 and 1954, the Supreme Court grappled with the subject of parallel conduct by oligopolists.<sup>5</sup> In the first two of these cases, *Interstate Circuit* and *American Tobacco*, the Supreme Court employed language implying that in some cases parallel conduct – standing alone – could be sufficient to prove conspiracy. With perhaps more interest in creating a bright line for lower courts and market participants than in addressing the potentially adverse economic impacts of “ordinary” oligopolistic behavior, the Supreme Court in *Theatre Enterprises* pulled back, decreeing that parallel action alone, even if undertaken with knowledge of the likely response of other market participants, does not violate Sherman Act Section 1.<sup>6</sup> As these cases

\* Richard Duncan is a partner in the Minneapolis office of Faegre & Benson LLP. Brian McCormac is an associate in the Des Moines office of Faegre & Benson LLP. An earlier version of this paper was presented at The Eighth Annual Sedona Conference on Antitrust Law & Litigation, Oct. 2006.

1 Compare Donald F. Turner, *The Definition of Agreement under the Sherman Act: Conscious Parallelism and Refusals to Deal*, 75 Harv. L. Rev. 655 (1962), with Richard A. Posner, *Oligopoly and the Antitrust Laws: A Suggested Approach*, 21 Stan. L. Rev. 1562 (1969).

Section 1 of the Sherman Act provides that “[e]very contract, combination . . . or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal.” 15 U.S.C. Section 1. Section 4 of the Clayton Act, 15 U.S.C. Section 15, provides a private right of action for injunctive relief and treble damages against violators of the Sherman Act.

2 Herbert Hovenkamp, *Federal Antitrust Policy*, Section 4.2 (1994).

3 *Id.*

4 *Id.*

5 See *Theatre Enters. v. Paramount Film Distrib. Corp.*, 346 U.S. 537 (1954); *American Tobacco Co. v. United States*, 328 U.S. 781 (1946); *Interstate Circuit, Inc. v. United States*, 306 U.S. 208 (1939).

6 *Theatre Enters.*, 346 U.S. at 541 (stating that “[c]ircumstantial evidence of consciously parallel behavior may have made heavy inroads into the traditional judicial attitude toward conspiracy; but ‘conscious parallelism’ has not read conspiracy out of the Sherman Act entirely”).

and their lower-court progeny illustrate, the federal courts have struggled to define the proof required to establish – and consequently the pleading required to allege – the existence of an agreement in parallel conduct cases.<sup>7</sup>

This term, in *Bell Atlantic Corp. v. Twombly*,<sup>8</sup> the Supreme Court returned to the subject of standards of pleading conspiracy in an oligopolistic setting – the market for local telephone and internet services. The Supreme Court granted certiorari in *Twombly* to review a decision of the United States Court of Appeals for the Second Circuit which held that a plaintiff asserting a conspiracy claim under Section 1 of the Sherman Act can survive a motion to dismiss by pleading only parallel conduct among the defendants. In *Twombly*, the plaintiffs brought a putative class action on behalf of themselves and all others who purchased local telephone or high-speed internet service in the continental United States on or after February 8, 1996, the effective date of the Telecommunications Act of 1996.<sup>9</sup> Plaintiffs' Amended Complaint alleged – wholly on information and belief – that defendants, four of the “Baby Bell” regional local telephone and internet companies, (a) conspired to prevent competitive entry into each of their local telephone and internet service markets and (b) agreed not to compete with each other and to allocate customers and markets.

The United States District Court for the Southern District of New York dismissed plaintiffs' Amended Complaint because plaintiffs alleged only that the defendants engaged in parallel conduct which was consistent with each of their individual economic interests. Because plaintiffs failed to plead any “plus factors” – facts that tend to exclude independent action as an explanation for defendants' parallel conduct – the district court found that the factual allegations of the Amended Complaint, even if true, failed to establish a cause of action.<sup>10</sup>

In an opinion authored by Judge Sack, a panel of the Second Circuit reversed, finding the pleading standard imposed by the district court inconsistent with the standards of Federal Rules of Civil Procedure 8 and 12.<sup>11</sup> The Second Circuit noted that while, in the absence of direct evidence of conspiracy, a plaintiff must show “plus factors” to prevail on a Section 1 claim at summary judgment and trial, a plaintiff need not allege “plus factors” to avoid dismissal at the pleading stage.<sup>12</sup> The court sympathized with the defendants' argument that allowing scant complaints to survive a motion to dismiss will open the doors to burdensome discovery and induce defendants to settle meritless claims to avoid the “colossal expense” of antitrust discovery, but declined to adopt what it characterized as a heightened pleading requirement in conspiracy cases under Section 1.<sup>13</sup> Instead, the court expressly invited Congress or the Supreme Court to “re-calibrate” the balance between the liberal pleading requirements of the Federal Rules and the realities of costly litigation.<sup>14</sup>

The Supreme Court accepted the invitation and reversed the Second Circuit, holding that a complaint alleging a Section 1 claim based on parallel conduct and a bare assertion of conspiracy will not survive a motion to dismiss.<sup>15</sup> However, the Supreme Court did not stop there, and proceeded to articulate a new “plausibility” standard for pleading Section 1 claims.<sup>16</sup> To survive a motion to dismiss under this standard, a Section 1 complaint must state “plausible grounds to infer an agreement” or, as the court explained, a complaint must contain “enough fact[s] to raise a reasonable expectation that discovery will reveal evidence of [an] illegal agreement.”<sup>17</sup> In articulating its new standard of “plausibility,” the court expressly rejected the 50-year-old formulation of Justice Black in *Conley v. Gibson* that a complaint may only be dismissed if a plaintiff “can prove no set of facts in support of his claim.”<sup>18</sup>

7 The Sedona Conference has also previously explored the issue of proof of agreement, though more in the context of summary judgment motions than standards in initial pleadings. See, e.g., Tyler A. Baker, *Proof of Conspiracy in Vertical and Horizontal Price-Fixing Cases: The Intersection of Law, Economics, Policy & Precedent*; William S. Comanor, *Demonstrating an Agreement Under the Sherman Act: Posner vs. Turner*, and Robert L. Hubbard, *Grants, Winks & Nods: What Meets the Agreement Element of a Section 1 Claim?*, all papers presented at The Fifth Annual Sedona Conference on Antitrust Law & Litigation, Oct. 2003.

8 *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 550 U.S. \_\_\_\_ (May 21, 2007).

9 Pub. L. No. 104-104, 110 Stat. 56 (codified at various sections of Titles 15 and 47 of the U.S. Code).

10 *Twombly v. Bell Atlantic Corp.*, 313 F. Supp. 2d 174, 189 (S.D.N.Y. 2003) (Gerald Lynch, J.).

11 The panel consisted of Judges Sack, Raggi, and Hall.

12 *Twombly v. Bell Atlantic Corp.*, 425 F.3d 99, 118 n.5 (2d Cir. 2005) (stating that “plaintiffs pursuing section 1 actions are not required to plead ‘plus factors’ to state a claim of conspiracy based on parallel anticompetitive conduct”).

13 *Id.* at 117.

14 *Id.*

15 *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 1961, 550 U.S. \_\_\_\_ (May 21, 2007).

16 *Id.* at 1965-66.

17 *Id.*

18 *Conley v. Gibson*, 355 U.S. 41, 45-46 (1957).

The Supreme Court clearly articulated its concerns with the skyrocketing cost of antitrust litigation, particularly during the discovery phase. Indeed, the court explained in detail its opinion on the negative aspects of large-scale discovery common (but not unique) to antitrust litigation.<sup>19</sup> The court emphasized that judges are unable to place meaningful controls on such discovery and that the financial costs of sweeping discovery “will push cost-conscious defendants to settle even anemic cases.”<sup>20</sup>

In antitrust law, it appears that *Twombly* has harmonized the pleading standard in Section 1 cases with the proof standard required to survive summary judgment under *Matsushita*.<sup>21</sup> The more onerous pleading standard announced in *Twombly* could increase the number of Section 1 cases resolved on motions to dismiss, and will certainly deter future plaintiffs from asserting Section 1 claims based on parallel conduct with no detailed facts indicating proof of agreement.

While *Twombly* need not necessarily apply outside the context of Section 1 cases, just as *Matsushita* has come to be generally relied on throughout the law of Rule 56 summary judgment, *Twombly*'s “plausibility” standard could soon come to dominate Rule 12 jurisprudence. In addition, the dicta from *Twombly* discussing the negative repercussions of large-scale discovery will undoubtedly find its way into countless briefs supporting motions for protective orders in complex cases. This article reviews the facts and law that brought the Supreme Court to reaffirm the principal of *Theatre Enterprises*, to reject *Conley v. Gibson*, and to set a new course in reviewing the adequacy of pleading under the Federal Rules of Civil Procedure.

## I. TWOMBLY'S ALLEGATIONS

On April 14, 2003, plaintiffs William Twombly and Lawrence Marcus filed an Amended Class Action Complaint against defendants Bell Atlantic Corporation (and its successor in-interest Verizon Communications, Inc.); BellSouth Corporation; Qwest Communications International, Inc.; and SBC Communications, Inc.<sup>22</sup> The Amended Complaint alleged that the defendants had conspired to prevent competitive entry into each of their local telephone and internet service markets and agreed not to compete with each other in violation of Section 1 of the Sherman Act. All of the substantive allegations contained in the Amended Complaint were based on “information and belief pursuant to the investigation of counsel.”<sup>23</sup>

The defendants were the four regional local telephone service providers that resulted from the 1984 divestiture of AT&T's local telephone business, commonly known as the “Baby Bells.”<sup>24</sup> Each of the defendants is a local exchange carrier, which provides customers with local telephone service.<sup>25</sup>

Until 1996, state regulatory authorities granted local exchange carriers exclusive franchises that allowed them to operate in their defined geographic areas without competition.<sup>26</sup> However, the Telecommunications Act of 1996 (the “Act”) attempted to stimulate competition in the local telephone markets in which local exchange carriers had enjoyed government-granted monopolies. Specifically, the Act required incumbent local exchange carriers that controlled access to the network infrastructure (such as defendants), to grant competitors access and connections to their lines and equipment on just, reasonable, and non-discriminatory terms.<sup>27</sup> The Act also required local exchange carriers that provided telephone exchange services to a defined area on the February 8, 1996 effective date of the Act to provide competitors with the same quality of service that they provide to themselves or their customers.<sup>28</sup>

19 *Bell Atlantic v. Twombly*, 127 S. Ct. at 1966-67.

20 *Id.* at 1967.

21 *Matsushita Elec. Indust. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986) (holding that a plaintiff must show evidence tending to exclude the possibility that the alleged conspirators acted independently to survive summary judgment under Section 1).

22 Mr. Twombly filed his initial Complaint on December 23, 2002, making virtually identical allegations to those contained in the Amended Complaint.

23 Amended Compl. at p. 1.

24 *Id.* at Paragraph 21. Initially, the divestiture of the Bell System by AT&T resulted in seven Regional Bell Operating Companies. However, by the time the Amended Complaint in Twombly was filed, later mergers and acquisitions had trimmed the number of Baby Bells to the four named defendants. *Id.* (During the pendency of the case, BellSouth acquired AT&T, and took the AT&T name for the new combined company.)

25 *Id.* at Paragraph 22.

26 *Id.* at Paragraph 23.

27 47 U.S.C. Section 251(c); Amended Compl. Paragraph 27.

28 47 U.S.C. Section 251(c)(2)(C); Amended Compl. Paragraph 28.

### A. *Twombly's* Conspiracy Allegations

The Amended Complaint recited a host of declarations and quotations gleaned from newspaper articles and statements by consumer advocacy groups claiming that the defendants “have refrained from engaging in meaningful head-to-head competition in each other’s markets.”<sup>29</sup> Plaintiffs asserted that this lack of competition “is strongly suggestive of conspiracy”<sup>30</sup> and that it was “especially unlikely” that neighboring defendants would have refrained from competing in each other’s adjacent geographic territories “[i]n the absence of an agreement not to compete.”<sup>31</sup> Finally, the Amended Complaint quoted a newspaper article reciting a statement from the CEO of Qwest that it would be fundamentally wrong to compete in SBC’s territory and that “it might be a good way to turn a quick dollar but that doesn’t make it right.”<sup>32</sup> The Amended Complaint summarized the defendants’ failure to compete with each other as indicative of their “strategy . . . to divide the country into local phone ‘fiefdoms,’ not to compete against each other, and to devote their collective efforts to ‘eliminating would-be competitors in local service.’”<sup>33</sup>

Plaintiffs alleged that the defendants “do indeed communicate amongst themselves” through various trade associations and that these associations provided them with the opportunity to coordinate their alleged conspiracy.<sup>34</sup> However, plaintiffs did not allege that defendants actually communicated in furtherance of their purported conspiracy.

Most notably, plaintiffs alleged that “[d]efendants have engaged in parallel conduct in order to prevent competition in their respective local telephone and/or high speed internet services markets.”<sup>35</sup> Specifically, plaintiffs claimed that defendants committed “one or more of the following wrongful acts in furtherance of a common anticompetitive objective:”

- failing to provide competitors with the same quality of service as provided to customers;
- failing to provide access to “operational support systems,” such as customer service records, on a nondiscriminatory basis;
- delaying the delivery of “unbundled elements” which precludes competitors from offering their customers service as attractive as that offered by defendants;
- billing their former customers after the customers converted service to a competitor, which resulted in double-billing and caused customer relations problems;
- providing interconnection between their networks and the networks of competitors at a lower quality than the interconnection that each defendant provided itself;
- refusing to sell competitors access to unbundled network components on just, reasonable, and nondiscriminatory terms;
- refusing to sell competitors local telephone and high-speed internet services at just, reasonable, and nondiscriminatory wholesale prices;
- refusing to allow competitors to connect to essential facilities, including local telephone lines and equipment, on just, reasonable, and nondiscriminatory terms;
- using “discriminatory and error-filled methods” to bill competitors, which discouraged competition “by making it virtually impossible for competitors to audit the bills they received from Defendants;”

<sup>29</sup> Amended Compl. at Paragraphs 37-39, 42, 44-45.

<sup>30</sup> *Id.* at Paragraph 40.

<sup>31</sup> *Id.* at Paragraph 41.

<sup>32</sup> *Id.* at Paragraph 42.

<sup>33</sup> *Id.* at Paragraph 44 (quoting *Illinois CLECS Assail Notebaert*, State Telephone Regulation Report, Comment, Vol. 20, No. 22 (Nov. 8, 2002)).

<sup>34</sup> *Id.* at Paragraph 46.

<sup>35</sup> *Id.* at Paragraph 47.

- imposing “slow and inaccurate manual order processing” that caused competitors to devote “significant time, effort, and expense” to correct problems and ensure that orders were processed correctly;
- using their monopoly power to gain a competitive advantage in the retail market for local telephone and high-speed internet services; and
- using their monopoly power and superior bargaining power to compel competitors to agree to unfair terms in exchange for access to local telephone networks.<sup>36</sup>

Plaintiffs provided maps showing in particular the balkanized service areas of defendant Verizon, which plaintiffs alleged should have induced competitive entry from other carriers.

## B. Market Structure Allegations

Plaintiffs alleged that the structure of the market for local telephone services facilitated the alleged conspiracy. Plaintiffs alleged that the four defendants controlled 90% or more of the local telephone market within the continental United States.<sup>37</sup> Thus, according to plaintiffs, “[e]laborate communications . . . would not have been necessary to enable Defendants to agree to allocate territories and to refrain from competing with one another.”<sup>38</sup>

Plaintiffs further alleged that the structure of the market for local telephone services facilitated a geographic market allocation scheme because the defendants would have quickly discovered if one of them began competing in the territory of another.<sup>39</sup> According to the Amended Complaint, this ease of detection made “a territorial allocation agreement among the Defendants more feasible, more readily enforceable, and more probable.”<sup>40</sup>

The Amended Complaint also alleged that had each defendant not illegally excluded competitors from their geographic territories, “the resulting competitive inroads into that Defendant’s territory would have revealed the degree to which competitive entry . . . would have been successful in other territories in the absence of such conduct.”<sup>41</sup> Plaintiffs also claimed that had a competitor made “substantial competitive inroads” into one of the defendants’ territories, it would have been more likely that other competitors would have made gains into the other defendants’ territories.<sup>42</sup> Thus, plaintiffs concluded that the defendants “had compelling common motivations to include in their unlawful horizontal agreement an agreement that each of them would engage in a course of concerted conduct calculated to prevent effective competition” in each of their allocated territories.<sup>43</sup>

## II. THE DISTRICT COURT’S DECISION

The defendants moved to dismiss the Amended Complaint on the ground that plaintiffs failed to allege facts sufficient to infer a conspiracy under Section 1 of the Sherman Act.<sup>44</sup> The district court granted defendants’ motion to dismiss, holding that the plaintiffs’ bare allegations of parallel conduct – which the court concluded coincided with each defendant’s individual economic interests – were insufficient, as a matter of law, to survive a motion to dismiss under Rule 12(b)(6), Fed. R. Civ. P.<sup>45</sup>

The court began its analysis by reciting the rule of *Theatre Enterprises* that parallel business behavior alone is not unlawful. The court also observed that although “parallel action is a common

<sup>36</sup> *Id.*

<sup>37</sup> *Id.* at Paragraph 48.

<sup>38</sup> *Id.*

<sup>39</sup> *Id.* at Paragraph 49.

<sup>40</sup> *Id.*

<sup>41</sup> *Id.* at Paragraph 50.

<sup>42</sup> *Id.*

<sup>43</sup> *Id.*

<sup>44</sup> *Twombly*, 313 F. Supp. 2d at 178. The district court did not address various other grounds asserted by defendants in support of their motions to dismiss. *Id.* at 178, 189 n.6.

<sup>45</sup> *Id.* at 189.

and often legitimate phenomenon, because similar market actors with similar information and economic interests will often reach the same business decisions,” parallel behavior can sometimes suggest that competing firms have entered into an illicit agreement.<sup>46</sup>

To determine whether parallel behavior is the product of legitimate market activity or an illegal conspiracy, the Second Circuit has required plaintiffs to show at least one “plus factor” that tends “to exclude independent self-interested conduct as an explanation for defendants’ parallel behavior.”<sup>47</sup> Plus factors include facts demonstrating that the parallel behavior was contrary to the economic interests of individual defendants or that a strong common motive to conspire existed among defendants. The district court acknowledged that the Second Circuit has only required plaintiffs to show plus factors at the summary judgment stage, but concluded that plaintiffs must allege plus factors to show that they have stated a claim for relief and that they are entitled to engage in discovery:

[P]laintiffs must always assert facts that, if true, support the existence of a conspiracy, such as motivation or conduct that lends itself to an inference of an agreement. In the context of parallel conduct allegations, simply stating that the defendants engaged in parallel conduct, and that this parallelism must have been due to an agreement, would be equivalent to a conclusory “bare bones” allegation of conspiracy. Viewed in this light, the plus factors are simply examples of allegations that are sufficiently suggestive of a conspiracy to warrant discovery.<sup>48</sup>

The court noted the tension between requiring plaintiffs to allege “specific facts suggesting a conspiracy,” and Rule 8, Fed. R. Civ. P., which requires only a “short and plain statement of the claim.”<sup>49</sup> However, the court explained that requiring a plaintiff to plead plus factors is consistent with the liberal requirements of Rule 8 for two reasons. First, parallel conduct often lawfully results from competitors with similar economic interests making decisions on the same information. Thus, allowing a plaintiff to open the door to discovery by pleading only allegations of parallel conduct “circumvents” both the foundational requirement of Section 1 that defendants actually engaged in a conspiracy and Rule 8’s requirement that complaints must state a claim for relief.<sup>50</sup> Second, pleading plus factors gives defendants notice of the plaintiff’s legal theories and “how and why” the defendants allegedly conspired.<sup>51</sup>

The court stated that a plaintiff asserting parallel conduct claims must allege facts that render the parallel conduct and state of the market “suspicious enough to suggest that defendants are acting pursuant to a mutual agreement rather than their own individual self-interest.”<sup>52</sup> These facts include “background propositions about how the market works when firms are competing, what it might look like if subject to an anti-competitive agreement, the economic interests of the market actors, and how those interests would cause them to act.”<sup>53</sup> The court conceded that these questions are similar to the factual issues in a summary judgment motion.<sup>54</sup>

The district court applied these standards to analyze the allegations of the Amended Complaint. The court noted that the geographic allocation alleged by plaintiffs did not show a conspiracy because the market had historically been insulated from open competition and the defendants’ geographic markets have been shaped by “decades of state and federal regulation.”<sup>55</sup> Accordingly, the operative questions were whether defendants’ alleged (a) efforts to keep competitors out of their territories and (b) refusal to compete in the territories of other defendants were products of the defendants’ individual economic interests or an illegal conspiracy.

---

<sup>46</sup> *Id.* at 179.

<sup>47</sup> *Id.*

<sup>48</sup> *Id.* at 180 (internal citations omitted).

<sup>49</sup> *Id.* at 180-182.

<sup>50</sup> *Id.* at 181.

<sup>51</sup> *Id.* at 181-82.

<sup>52</sup> *Id.* at 182.

<sup>53</sup> *Id.*

<sup>54</sup> *Id.*

<sup>55</sup> *Id.* at 183.



The court assessed plaintiffs' allegations that defendants conspired to prevent competitive entry into their territories and burdened competitors after they entered defendants' territories. The plaintiffs failed to allege any reasons to believe that defendants would have acted differently – as by welcoming competition into their territory – absent the illegal agreement alleged by plaintiffs. The court found that even if defendants discouraged competitors from entering their respective territories and made it difficult for competitors to survive after they did so, such actions were wholly consistent with each defendant's individual economic interests.<sup>56</sup> Since the plaintiffs had not alleged any basis to infer that each defendant's behavior was attributable to anything beyond pursuing its individual economic interests, the court dismissed plaintiffs' claim that the defendants illegally agreed to suppress competition in their territories.<sup>57</sup>

The court then reviewed plaintiffs' claim that each defendant agreed not to compete in the other defendants' territories and acknowledged that this theory "present[ed] a closer question."<sup>58</sup> The court questioned plaintiffs' theory that defendants' refusal to compete with each other was against their individual economic interests because defendants could not be expected to enter new markets unless they believed that doing so would be profitable.<sup>59</sup>

The court accepted defendants' arguments that a defendant seeking to compete in another defendant's territory would necessarily be forced to adopt a vastly different business model than it employs in its own territory. Indeed, each defendant, as the incumbent local exchange carrier ("ILEC") in its territory, controls and maintains the telecommunications network in its territory and determines the prices at which it offers customers (including competitive local exchange carriers, or "CLECs," who buy network time from incumbent carriers at wholesale for resale to retail customers) access to the network. If a defendant wishes to compete in the territory of another ILEC, it must obtain access to the network on the best terms it can negotiate from that ILEC. Then, the potential competitor must purchase access to the ILEC's network at wholesale and repackage the network time and sell it at a profit to retail customers in that territory.<sup>60</sup>

To the extent each defendant enjoys name recognition and goodwill in its territory, these intangible assets purportedly have very limited value outside of the defendant's territory because each defendant's ability to compete outside its territory depends heavily on its ability to obtain network access on favorable terms from the ILEC for that territory. Thus, each defendant cannot easily use its intangible assets to influence the market for local telephone and internet services outside of its territory. Accordingly, the district court hypothesized, each defendant would lose the inherent advantages of controlling the network infrastructure if it chose to enter markets outside of its territory.<sup>61</sup>

In addition, the court rejected plaintiffs' allegations that defendants operating in geographic proximity have competitive advantages in the territories of neighboring ILECs.<sup>62</sup> The court reemphasized that being an incumbent carrier and controlling the telecommunications network is a vastly different business than being a competing carrier and having to buy access to the network. According to the court, given the purchase and resale structure created by the Telecommunications Act, it would be "impossible" or "undesirable" to build a competing network (or to expand the existing network) into a competing territory.<sup>63</sup> Indeed, a defendant seeking to compete in an adjacent territory could not "leverage geographical proximity into independence, and therefore is not materially different from a . . . [competing local exchange carrier] without a nearby territory of its own."<sup>64</sup> Therefore, if a defendant sought to compete in another defendant's territory, it would face the same market challenges and entry barriers as any other local exchange carrier.<sup>65</sup> The court mentioned plaintiffs' allegations that each defendant has successfully thwarted competitive entry into its territory

---

<sup>56</sup> *Id.*

<sup>57</sup> *Id.* at 184, 189.

<sup>58</sup> *Id.* at 184.

<sup>59</sup> *Id.* at 185.

<sup>60</sup> *Id.* at 185-186.

<sup>61</sup> *Id.* at 186.

<sup>62</sup> *Id.* at 185-189.

<sup>63</sup> *Id.* at 186.

<sup>64</sup> *Id.* at 186-87.

<sup>65</sup> *Id.* at 187.

and stated that plaintiffs failed to plead any facts to create doubt about each defendant's unilateral ability to frustrate the competitive entry into its territory by another defendant.<sup>66</sup>

The court held that plaintiffs failed to generate an inference that the defendants' actions were the product of an illegal agreement because the Amended Complaint stated no basis to believe that any defendant would perceive an economic interest in competing in another defendant's territory.<sup>67</sup> Further, the court ruled that plaintiffs did not sufficiently allege that defendants had a common motive to conspire.<sup>68</sup> Thus, the court held that the plaintiffs alleged "no reason to believe that defendants' parallel conduct was reflective of any agreement" and concluded that plaintiffs failed to state a cause of action under Section 1 of the Sherman Act because the Amended Complaint "alleges nothing more than parallel conduct that appears to accord with the individual economic interests of the alleged conspirators."<sup>69</sup>

As the cited excerpts show, the district court engaged in what looks like fact-finding. The level of analysis is unusual in the context of a motion to dismiss and appears more akin to the summary judgment analysis mandated by *Matsushita*.

### III. THE SECOND CIRCUIT'S OPINION

The Second Circuit reversed, holding that the district court applied an incorrect standard for evaluating a complaint challenged on a motion to dismiss.<sup>70</sup> The court reiterated that no heightened pleading standards apply in antitrust cases.<sup>71</sup> Instead, the liberal standard of Fed. R. Civ. P. 8 applies to conspiracy claims, such as those asserted by plaintiffs.<sup>72</sup> Rule 8 provides that a plaintiff must allege a "short and plain statement of the claim showing that the pleader is entitled to relief." The appellate court noted that while Rule 8 only requires a complaint to contain "minimal" facts in support of its claims, a complaint having "nonexistent" facts should not survive a motion to dismiss.<sup>73</sup> Thus, a "bare bones" complaint lacking any factual predicate for conspiracy claims should be dismissed, as should a complaint that alleges an "implausible" conspiracy on the facts pleaded or that is premised on "unlikely speculations."<sup>74</sup>

The disagreement between the district court and the Second Circuit centered around the factual allegations necessary in a conspiracy case to survive a motion to dismiss. The district court's holding would require plaintiffs alleging an antitrust conspiracy based on parallel conduct to allege "plus factors" – facts which tend to exclude independent self-interested conduct as an explanation for the defendants' parallel behavior.<sup>75</sup> Otherwise, a plaintiff could survive a motion to dismiss by alleging facts that, even if true, fail to establish liability.

The Second Circuit acknowledged that a plaintiff in a parallel conduct case must prove the existence of "plus factors" to avoid summary judgment.<sup>76</sup> However, in the context of a motion to dismiss, a plaintiff must allege only the existence of a conspiracy and "a sufficient supporting factual predicate" for the conspiracy allegations.<sup>77</sup> According to the Second Circuit, the factual predicate for conspiracy claims need only "*include conspiracy among the realm of plausible possibilities in order to survive a motion to dismiss.*"<sup>78</sup> Thus, under the Second Circuit's test, a plaintiff may – but is not required to – plead plus factors in support of a conspiracy claim based on parallel conduct. According to the Second Circuit, in order to dismiss a conspiracy claim based on parallel conduct, "*a court would have to conclude that there is no set of facts that would permit a plaintiff to demonstrate that the particular parallelism asserted was the product of collusion rather than coincidence.*"<sup>79</sup>

66 *Id.* at 189.

67 *Id.* at 187.

68 *Id.* at 188.

69 *Id.* at 189.

70 *Twombly*, 425 F.3d at 102.

71 *Id.* at 107-09.

72 *Id.* at 108-09.

73 *Id.* at 109.

74 *Id.* at 109-111.

75 *Twombly*, 313 F. Supp. 2d at 179-80.

76 *Twombly*, 425 F.3d at 113-14.

77 *Id.* at 114.

78 *Id.* (emphasis added).

79 *Id.* (emphasis added).

Applying this test, the Second Circuit held that the allegations of the Amended Complaint gave defendants “fair notice” of the claims asserted against them and the grounds upon which those claims were based.<sup>80</sup> The court acknowledged that while the Amended Complaint failed to “identify specific instances of conspiratorial conduct or communications,” it did state “the temporal and geographic parameters of the alleged illegal activity and the identities of the alleged key participants.”<sup>81</sup> The court recognized that plaintiffs alleged that defendants have “frequent opportunities to organize and conduct their conspiracy through industry organizations.”<sup>82</sup> The court further opined that the Amended Complaint alleged that defendants shared a “common incentive” to conspire to prevent any competitive local exchange carrier from tasting success in one of the defendant’s territories, which could embolden other local exchange carriers to enter into the defendants’ territories.<sup>83</sup> Of course, similar allegations could be made in the abstract about virtually any industry characterized by oligopoly.

While the First, Sixth and Tenth Circuits had concluded that a complaint alleging a Section 1 claim based on parallel conduct had to allege “plus factors” in order to survive a motion to dismiss,<sup>84</sup> the Second Circuit was not alone in holding that a plaintiff need not plead “plus factors” to state such a claim. The Third Circuit said as much in 2004 in *Lum v. Bank of America*,<sup>85</sup> in a unanimous decision in which then-Judge Samuel Alito joined. *Lum* based its holding on the 1977 case of *Bogosian v. Gulf Oil Corp.*<sup>86</sup> *Bogosian*, in turn, held that the existence of collusion in a highly concentrated market characterized by interdependent decision-making among its participants was too complex to be decided based on the complaint alone:

If these theories are to be tested, it should be done on a fully developed factual record which probes the conflicting economic facts on which they are premised. The complaint is much too blunt an instrument with which to forge fundamental policies regarding the meaning of competition in concentrated industries.<sup>87</sup>

*Bogosian* thus implied that a complaint alleging conscious parallelism can state a Section 1 claim. In *Twombly* the Second Circuit cited *Lum* only in passing<sup>88</sup> and *Bogosian* not at all.

The Court of Appeals also noted the district court’s “well-founded concern” that allowing conspiracy cases to proceed solely on allegations of parallel conduct would “condemn defendants to potentially limitless ‘fishing expeditions’ – discovery pursued just ‘in case anything turn[s] up’ – in hopes, perhaps, of a favorable settlement.”<sup>89</sup> In addition, the court recognized the other concerns articulated by the district court that, in its view, made requiring plaintiffs to plead plus factors “sensible”: (1) the antitrust laws do not prohibit parallel conduct; and (2) a complaint should give defendants notice of plaintiffs’ theory of conspiracy.<sup>90</sup> On appeal, the defendants had argued that if plaintiffs need not plead plus factors, then “any claim asserting parallel conduct [would] survive a motion to dismiss.”<sup>91</sup>

In response to these concerns, the Second Circuit merely stated:

80 *Id.* at 118-19.

81 *Id.* at 117.

82 *Id.* at 118.

83 *Id.*

84 See *DM Research, Inc. v. College of American Pathologists*, 170 F.3d 53 (1st Cir. 1999); *NHL Players Association v. Plymouth Whalers Hockey Club*, 419 F.3d 462 (6th Cir. 2005); *Cayman Exploration Corp. v. United Gas Pipe Line Co.*, 873 F.2d 1357 (10th Cir. 1989).

85 *Lum v. Bank of America*, 361 F.3d 217, 230 (3d Cir. 2004), cert. denied, 543 U.S. 918 (2004) (“Since conscious parallelism is an evidentiary rule that relates to how a plaintiff may prove the existence of an agreement, a plaintiff need not allege the existence of these plus factors in order to plead an antitrust cause of action.”).

86 *Bogosian v. Gulf Oil Corp.*, 561 F.2d 434 (3d Cir. 1977).

87 *Id.* at 447 (citing Posner, *supra* note 1, Turner, *supra* note 1, L. Sullivan, Law of Antitrust Section 122 (1977), and *White Motor Co. v. United States*, 372 U.S. 253 (1963)). The explicit judicial grappling with the theoretical and practical problems for antitrust law posed by oligopolies seen in *Bogosian* was not presented to the Supreme Court in the parties’ briefs in *Twombly* (nor reflected in the court’s ultimate decision). Instead the parties’ arguments focused on less relevant market structures (*i.e.*, perfect competition) and less relevant areas of law (*i.e.*, the employment law decision in *Swierkiewicz v. Sorema N.A.*, 534 U.S. 506 (2002)).

88 *Twombly*, 425 F.3d at 114 n.9.

89 *Id.* at 115-16 (alteration in original).

90 *Id.* at 116.

91 *Id.* (alteration in original).

We are not unsympathetic to these concerns, but we find the arguments based on them ultimately unconvincing. At the pleading stage, we are concerned only with whether the defendants have “fair notice” of the claim, and the conspiracy that is alleged as part of the claim, against them – that is, enough to “enable [the defendants] to[, *inter alia*,] answer and prepare for trial,” not with whether the conspiracy can be established at trial.<sup>92</sup>

In support of its holding, the court reemphasized that neither the Federal Rules of Civil Procedure nor Supreme Court precedent place a heightened pleading burden on plaintiffs alleging an antitrust conspiracy.

In an unusual passage undercutting the public policy of the lenient federal pleading rules on which it based its decision, the Second Circuit then invited Congress or the Supreme Court to re-evaluate the proper pleading standard:

We are mindful that a balance is being struck here, that on one side of that balance is the sometimes colossal expense of undergoing discovery, that such costs themselves likely lead defendants to pay plaintiffs to settle what would ultimately be shown to be meritless claims, that the success of such meritless claims encourages others to be brought, and that the overall result may well be a burden on the courts and a deleterious effect on the manner in which and efficiency with which business is conducted. If that balance is to be re-calibrated, however, it is Congress or the Supreme Court that must do so.<sup>93</sup>

The Second Circuit’s comments on the purported costs of discovery and prevalence of nuisance settlements were asserted without specific support. Nonetheless, the court essentially invited review – and reversal – by the Supreme Court.

#### IV. SUPREME COURT BRIEFING

Defendants filed a petition for a writ of certiorari to review the Second Circuit’s opinion, asserting several grounds in support of review. First, they argued that because unilateral parallel conduct is not unlawful, a complaint alleging a Section 1 violation based on parallel conduct must allege a factual basis to support an inference of conspiracy. Second, defendants claimed that the Second Circuit’s decision conflicted with the law of the Tenth, First, and Sixth Circuits, which require plaintiffs alleging a conspiracy under Section 1 to plead facts in support of their conspiracy claims. Third, defendants made policy arguments concerning the potential explosion of frivolous litigation, fishing expeditions, and ransom settlements, and the corresponding burdens on courts and businesses.

Several entities filed *amicus curiae* briefs urging the Supreme Court to reverse the Second Circuit. Of particular interest are the positions articulated by a group of economists and by the United States.

The economists argued that parallel behavior is a common feature in competitive markets. Firms should be expected to react similarly to market changes, such as demand and supply shifts, changes in consumer preferences, and technological changes.<sup>94</sup> The economists claimed that allowing plaintiffs to conduct discovery solely on pleadings alleging parallel conduct could discourage firms from “responding efficiently to market signals in ways that might seem to be mimicking the conduct of rivals if they faced the prospect that actions similar to those taken by competitors could support at least the discovery phase of antitrust claims.”<sup>95</sup> Thus, the economists argued that the Second Circuit’s decision and the corresponding risk of burdensome litigation would cloud the ability of firms to make legitimate business decisions taken by all firms, such as pricing, market entry and withdrawal, and

<sup>92</sup> *Id.* (alterations in original; internal citations omitted).

<sup>93</sup> *Id.* at 117.

<sup>94</sup> Brief of Amici Curiae Economists in Support of Petitioners, *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 550 U.S. \_\_\_\_ (No. 05-1126), 2006 WL 2506633 (Aug. 25, 2006).

<sup>95</sup> *Id.* at \*3.

marketing strategies. In addition, the economists argued that the inability of courts to dismiss meritless claims early in the litigation process will create negative repercussions including imposing unnecessary direct costs on firms and deadweight costs on the economy.

The economists' brief did not address the actual market structure at issue in *Twombly*, an oligopoly (or perhaps a series of interrelated local monopolies), and the serious question of regulating – or not regulating – the adverse impacts to competition that can arise from that market structure. The economists posit only a caricature of the issues raised by plaintiffs' Amended Complaint, asserting that, if affirmed, the Second Circuit's decision would allow a price-fixing class action "against virtually any set of two or more competing firms in the economy" to proceed to discovery.<sup>96</sup> Surely such cases would fall prey even to the Second Circuit's vaguely-stated "plausibility" standard. The economists failed to engage the Supreme Court in the proper application of Section 1 to the specific market structure at issue in *Twombly* – the distressingly resilient oligopoly in local telephone and internet service.

The United States claimed that it has a substantial interest in this case because of the negative effects of meritless antitrust suits, which include creating economic inefficiencies, chilling pro-competitive conduct, and acting as a drain on the economy because of the significant costs incurred by defendants, whether in the form legal fees and expenses or *in terrorem* settlement payments to avoid burdensome discovery.<sup>97</sup> The United States argued that the Second Circuit applied the wrong standard in analyzing the Amended Complaint in light of the defendants' motion to dismiss.<sup>98</sup> The United States contended that parallel activity "is a hallmark of competitive markets" because over time, firms tend to adopt similar business practices and strategies that have proven to be reliable and efficient.<sup>99</sup> Thus, the United States argued, parallel action, standing alone, cannot give rise to an antitrust claim. Further, inferring a conspiracy on the basis of parallel inaction is even more problematic because "the number of territories a business does not enter or products it does not offer is virtually infinite."<sup>100</sup>

The United States suggested the following standard of review on motions to dismiss:

[T]he complaint [must] give rise to a reasonably grounded expectation that discovery will reveal relevant evidence to support [a] Sherman Act claim, and . . . provide the requisite fair notice to [defendants] of the basis for that claim.<sup>101</sup>

The United States expressly stated that its proposed standard of review does not go so far as to require a Section 1 complaint to allege facts "that would be sufficient to defeat summary judgment under the standard articulated in *Matsushita Electric*."<sup>102</sup> The United States acknowledged three times that *Twombly* presents a "close question," but ultimately concluded that the Amended Complaint fails to satisfy the requirements of Rule 8.<sup>103</sup>

The Government's *amicus* brief illustrates the fundamental difficulty in assessing the sufficiency of the allegations in *Twombly*. The Government conceded that the case is a close one; proffered a standard of review that does not expressly require pleading of plus factors and vaguely invited courts to evaluate the "reasonableness" of a plaintiff's allegations in the context of the plainly non-competitive market at issue, all of which goes beyond traditional Rule 12 analysis such as that set forth in *Conley v. Gibson*.

96 *Id.* at \*10.

97 Brief for the United States as Amicus Curiae Supporting Petitioners, *Bell Atlantic Corp., v. Twombly*, 127 S. Ct. 1955, 550 U.S. \_\_\_\_ (No. 05-1126), 2006 WL 2482696 (Aug. 25, 2006), at \*1.

98 *Id.* at \*9 (stating that the legal standard adopted by the Second Circuit "is wrong and has the potential to chill substantial economic activity that is both efficient and innocuous from the standpoint of the antitrust laws").

99 *Id.* at \*20.

100 *Id.* at \*21.

101 *Id.* at \*26.

102 *Id.* at \*22-\*23 (citing *Matsushita*, 475 U.S. at 588).

103 *Id.* at \*8, \*10, \*26.

## V. THE SUPREME COURT'S OPINION

The Supreme Court stated the issue and holding in its *Twombly* decision as follows: “[t]he question in this putative class action is whether a Section 1 complaint can survive a motion to dismiss when it alleges that major telecommunications providers engaged in certain parallel conduct unfavorable to competition, absent some factual context suggesting agreement, as distinct from identical, independent action. We hold that such a complaint should be dismissed.”<sup>104</sup> Despite this basic characterization, the Supreme Court’s seven-member majority opinion also overrules a long-standing and oft-cited case interpreting Rules 8 and 12 of the Federal Rules of Civil Procedure, arguably rewrites the notice pleading standard embodied in those rules, and deplors the costs of discovery in complex cases.

### A. “Plausibility” Pushes Aside *Conley v. Gibson*

The Supreme Court began its analysis with the familiar statement that the crucial question in Section 1 cases is whether the alleged anticompetitive conduct “stems from independent decision or from an agreement, tacit or express.”<sup>105</sup> The court then reiterated that parallel conduct, standing alone, cannot establish a successful Section 1 claim because it is “consistent with conspiracy, but just as much in line with a wide swath of rational and competitive business strategy unilaterally prompted by common perceptions of the market.”<sup>106</sup> In so holding, the court relied on Professor Turner’s 1962 article. Thus, and without substantive analysis of the consequences of its holding for competitively dysfunctional markets such as local telephone service, the court reaffirmed the rule of *Theatre Enterprises*.

The Supreme Court then launched into a discussion of the pleading requirements of Rule 8 and their relationship to the motion to dismiss standard set forth in Rule 12(b)(6).<sup>108</sup> Although Rule 8 does not explicitly require the pleading of “facts,” the court noted that “a plaintiff’s obligation to provide the ‘grounds’ of his ‘entitlement to relief’ requires more than labels and conclusions. . . . Factual allegations must be enough to raise a right to relief above the speculative level.”<sup>109</sup> Indeed, the court questioned whether it would be possible for a plaintiff to satisfy Rule 8’s pleading requirements of providing “‘fair notice’ of the nature of the claim” and “the ‘grounds’ on which the claims rests” without some factual allegations.<sup>110</sup>

The Supreme Court applied these “general standards” and developed a “plausibility” test: in order to state a claim under Section 1, a plaintiff must plead “enough factual matter to suggest that an agreement was made.”<sup>111</sup> The court insisted that requiring a plaintiff to plead plausible factual grounds to infer the existence of an illegal agreement “does not impose a probability requirement at the pleading stage.” Rather, the court maintained that its plausibility standard “simply calls for enough fact [sic] to raise a reasonable expectation that discovery will reveal evidence of [an] illegal agreement.”<sup>112</sup> The court explained that requiring allegations plausibly suggesting – and not simply consistent with – the existence of an agreement “reflects the threshold requirement of Rule 8(a)(2) that the ‘plain statement’ possess enough heft to ‘show that the pleader is entitled to relief.’”<sup>113</sup> Bare allegations of parallel conduct “get[ ] the complaint close to stating a claim, but without some further factual enhancement it stops short of the line between possibility and plausibility of ‘entitlement to relief.’”<sup>114</sup>

The Supreme Court’s “plausibility” test appears grounded in a very strong aversion to allowing potentially “false positive” antitrust claims to proceed to the discovery stage. In explaining the justifications for its plausibility standard, the court relied upon Seventh Circuit Judges Richard

104 *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 1961, 550 U.S. \_\_\_\_ (May 21, 2007).

105 *Id.* at 1964 (quoting *Theatre Enters., Inc. v. Paramount Film Distrib. Corp.*, 346 U.S. 537, 541 (1954)) (internal quotations and alteration omitted).

106 *Id.*

107 *Id.*

108 *Id.* at 1964-65.

109 *Id.* (internal alterations omitted).

110 *Id.* at 1965, n.3.

111 *Id.* at 1965 (internal parenthetical omitted).

112 *Id.*

113 *Id.* at 1966 (quoting Fed. R. Civ. P. 8(a)(2)) (alteration omitted).

114 *Id.* (alteration omitted).

Posner and Frank Easterbrook, who have assailed the costly and burdensome nature of class action and antitrust discovery and the potential for *in terrorem* settlements putatively connected with such discovery (even though the court provides no empirical evidence of such settlements).<sup>115</sup> Indeed, as the court noted, “it is one thing to be cautious before dismissing an antitrust complaint in advance of discovery but quite another to forget that proceeding to antitrust discovery can be expensive.”<sup>116</sup> Quoting *Associated General Contractors*, the court reiterated that “a district court must retain the power to insist upon some specificity in pleading before allowing a potentially massive factual controversy to proceed.”<sup>117</sup> The court noted that the potential expense of discovery in *Twombly* could be staggering because the putative plaintiff class includes at least 90 percent of local phone or internet subscribers in the continental United States, the defendants have thousands of employees “generating reams and gigabytes of business records”, and the alleged antitrust violations span a seven-year period.<sup>118</sup>

The Supreme Court then found that district courts are ill-suited to weed out claims that are “just shy of a plausible entitlement to relief” because of the practical problems district courts face in managing discovery and the common belief that “the success of judicial supervision in checking discovery abuse has been on the modest side.”<sup>119</sup> Although unmeritorious claims may be disposed of on summary judgment, the court observed that the timing of a summary judgment motion would not avoid the expense and burdens of discovery.<sup>120</sup> Thus, the court concluded that requiring allegations that provide plausible grounds for inferring the existence of a conspiracy is the only probable way to avoid the potentially colossal expense of discovery in cases with no reasonable expectation that discovery will reveal evidence of an illegal agreement.<sup>121</sup>

Next, and surprisingly, the Supreme Court revisited *Conley v. Gibson*, the 1957 case containing the oft-cited language that “a complaint should not be dismissed for failure to state a claim unless it appears beyond doubt that the plaintiff can prove no set of facts in support of his claim which would entitle him to relief.”<sup>122</sup> The court rejected *Conley’s* “no set of facts” language because it could be read in isolation to allow a “wholly conclusory statement of claim” to survive a motion to dismiss unless the pleadings show the factual impossibility of the claim on their face.<sup>123</sup> Then, the court canvassed judicial and academic authority to underpin its rejection of *Conley* and concluded that “*Conley’s* ‘no set of facts’ language has been questioned, criticized, and explained away long enough.”<sup>124</sup> Thus, the court expressly overruled *Conley* to the extent its “no set of facts” language guided pleading practice: “after puzzling the profession for 50 years, this famous observation has earned its retirement.”<sup>125</sup>

## B. *Twombly* Failed to State a Plausible Claim under Section 1

Turning to the case before it, the Supreme Court agreed with the district court that plaintiff *Twombly’s* allegations failed to state a plausible claim of conspiracy.<sup>126</sup> Although the Supreme Court recognized that some of the paragraphs of the Amended Complaint speak of or refer to an agreement, it dismissed these allegations as mere legal conclusions.<sup>127</sup>

The Supreme Court then concluded that none of the specific claims pleaded by the plaintiffs state a cause of action because these claims do not allege a plausible suggestion of a conspiracy.<sup>128</sup> Instead, the court asserted, each instance of anticompetitive activity alleged by the plaintiffs was wholly consistent with legitimate unilateral economic activity. For example, plaintiffs

115 *Id.* at 1966-67 (citing *Asabi Glass Co. v. Pentech Pharmaceuticals, Inc.*, 289 F. Supp. 2d 986, 995 (N.D. Ill. 2003) (Posner, J., sitting by designation), and Frank Easterbrook, *Discovery as Abuse*, 69 B.U. L. Rev. 635 (1989)).

116 *Id.* at 1966-67 (internal citation omitted).

117 *Id.* at 1967 (quoting *Associated Gen. Contractors of Cal., Inc. v. Carpenters*, 459 U.S. 519, 528 n.17 (1983), and citing *Car Carriers, Inc. v. Ford Motor Co.*, 745 F.2d 1101, 1106 (7th Cir. 1984)).

118 *Id.*

119 *Id.*

120 *Id.*

121 *Id.*

122 *Id.* at 1968 (quoting *Conley v. Gibson*, 355 U.S. 41, 45-46 (1957)).

123 *Id.*

124 *Id.* at 1969.

125 *Id.*

126 *Id.* at 1970.

127 *Id.*

128 *Id.* at 1970-71.

claimed that the defendants suppressed competitive entry into the defendants' incumbent markets and that the defendants coordinated these efforts. The court opined that plaintiffs failed to plead any facts that would show that these actions by defendants, even if true, were anything other than "the natural, unilateral reaction of each [defendant] intent on keeping its regional dominance."<sup>129</sup> Indeed, the court quipped that "if alleging parallel decisions to resist competition were enough to imply an antitrust conspiracy, pleading a Section 1 violation against almost any group of competing businesses would be a sure thing."<sup>130</sup>

Plaintiffs also alleged conspiracy based on the failure of the defendants to begin competing in their neighbors' territory in the wake of the deregulation spawned by the 1996 Telecommunications Act.<sup>131</sup> The court made short work of this claim as well, noting that a natural explanation for the alleged noncompetition could be that each of the defendants was sitting tight and enjoying the vestigial effects of the regulated world of communications while expecting that their competitors would act the same way.<sup>132</sup> In addition, the court noted that "firms do not expand without limit and none of them enters every market that an outside observer might regard as profitable, or even a small portion of such markets,"<sup>133</sup> and plaintiffs failed to allege that the nonexistent inter-territorial competition would have been more lucrative than any other opportunities the defendants actually pursued.<sup>134</sup>

The Supreme Court concluded by asserting that its plausibility test is neither a heightened pleading standard, nor an expansion of the scope of Rule 9, which requires certain claims to be pleaded with greater particularity.<sup>135</sup> Thus, the court was able to square this decision with its most recent decision in the Rule 12 context, *Swierkiewicz v. Sorema N.A.*, 534 U.S. 506 (2002), which reversed a dismissal of an employment law complaint as having required too much specificity of the plaintiff. Instead, the court states in *Twombly* that it requires no "heightened fact pleading of specifics, but only enough facts to state a claim to relief that is plausible on its face."<sup>136</sup> The court's concern was not that the allegations of the Amended Complaint were not pleaded with sufficient particularity, but rather that the Amended Complaint "failed *in toto* to render plaintiffs' entitlement to relief plausible."<sup>137</sup> Thus, the court concluded, "because the plaintiffs have not nudged their claims across the line from conceivable to plausible, their complaint must be dismissed."<sup>138</sup>

Notwithstanding these explanations, the Supreme Court's new "plausibility" standard suffers from inherent indeterminacy, as the history of *Twombly* shows. The Second Circuit had *upheld* the sufficiency of the Amended Complaint because conspiracy was a "plausible" conclusion that could be drawn from the facts alleged.<sup>139</sup> The Supreme Court held that the Amended Complaint's allegations of conspiracy are *not* "plausible," and dismissed the case on that basis.<sup>140</sup> "Plausibility" is thus from its birth a legal standard that is highly subjective, much more so than *Conley v. Gibson's* "no set of facts" standard, which most courts have applied consistently to deny most motions to dismiss.

### C. The Dissent

Justices Stevens and Ginsburg dissented, assailing the "plausibility" standard articulated by the majority. The dissent begins by contending that the issue framed by the majority, *i.e.*, whether allegations of parallel conduct state a claim for relief under Section 1, can be answered in the negative based on the 50-year-old *Theatre Enterprises* case.<sup>141</sup> The dissent notes that the parallel conduct alleged in the Amended Complaint, which the defendants assert is consistent with the absence of an illegal

129 *Id.* at 1971.

130 *Id.*

131 *Id.* at 1972.

132 *Id.*

133 *Id.* at 1973 (quoting P. Areeda & H. Hovenkamp, Antitrust Law Paragraph 307d, at 155 (Supp. 2006)).

134 *Id.*

135 *Id.* at 1973-1974.

136 *Id.* at 1974. When Congress has specified a heightened pleading standard for particular claims, the current Supreme Court has been willing to apply it strictly. *E.g.*, *Tellabs, Inc. v. Makor Issues and Rights, Ltd.*, 127 S. Ct. 2499 (2007) (applying heightened pleading standard under Private Securities Litigation Reform Act of 1995, 109 Stat. 737).

137 *Id.* at 1973, n.14.

138 *Id.* at 1974.

139 *Twombly*, 425 F.3d at 114.

140 *Bell Atlantic v. Twombly*, 127 S. Ct. at 1974.

141 *Id.* (Stevens, J. dissenting) (citing *Theatre Enters., Inc. v. Paramount Film Distrib. Corp.*, 346 U.S. 537 (1954)).



agreement, is also consistent with the presence of an illegal agreement. Thus, the dissent maintains that “the Federal Rules of Civil Procedure, our longstanding precedent, and sound practice . . . at least require some sort of response from petitioners before dismissing the case.”<sup>142</sup>

The dissent identified two “practical concerns” that presumably explain the court’s “dramatic departure from settled procedural law”: the expense of antitrust litigation and the risk that a jury may use evidence of parallel conduct to determine the existence of an agreement, when, in fact, the parties were acting alone but made similar business decisions.<sup>143</sup> The dissent opines that these concerns do not justify dismissal, but rather call for “careful case management, including strict control of discovery, careful scrutiny of evidence at the summary judgment stage, and lucid instructions to juries.”<sup>144</sup>

The dissent reviewed the evolution from code pleading to the notice pleading standard presently embodied in the Federal Rules of Civil Procedure. After noting the historical difficulties caused by requiring the pleading of “facts,” the dissent concluded that “Rule 8 was directly responsive to this difficulty” and “[i]ts drafters intentionally avoided any reference to ‘facts’ or ‘evidence’ or ‘conclusions.’”<sup>145</sup>

The dissent defended *Conley v. Gibson*, and interpreted that decision’s “no set of facts” standard to allow dismissal “only when proceeding to discovery or beyond would be futile.”<sup>146</sup> The dissent took issue with the majority’s criticism of *Conley* and its statement that the “no set of facts” phrase has been “questioned, criticized, and explained away long enough” and “has puzzled the profession for 50 years.” The dissent noted that *Conley* has been cited by the Supreme Court in a dozen majority opinions and that 26 states and the District of Columbia use the “no set of facts” test to evaluate motions to dismiss.<sup>147</sup>

After reviewing a litany of cases discussing the liberality of federal pleading standards, the dissent concludes that despite its statements to the contrary, the majority created a heightened pleading standard that is “irreconcilable with Rule 8 and with our governing precedents.”<sup>148</sup> Indeed, according to the dissent, “[e]verything today’s majority says would . . . make sense if it were ruling on a Rule 56 motion for summary judgment . . . [b]ut it should go without saying . . . that a heightened production burden at the summary judgment stage does not translate into a heightened pleading burden at the complaint stage.”<sup>149</sup> The dissent cites the availability of treble damages and attorneys’ fees as evidence of congressional intent to encourage private enforcement of the antitrust laws, and admonishes, “[i]t is therefore more, not less, important in antitrust cases to resist the urge to engage in armchair economics at the pleading stage.”<sup>150</sup>

The dissent then reviewed the allegations of the Amended Complaint and concluded that dismissal was not warranted under existing pleading rules. Finally, Justice Stevens (Justice Ginsburg declined to join this Part) asked the pivotal question:

Whether the Court’s actions will benefit only defendants in antitrust treble-damages cases, or whether its test for the sufficiency of a complaint will inure to the benefit of all civil defendants, is a question that the future will answer. But that the Court has announced a significant new rule that does not even purport to respond to any congressional command is glaringly obvious.<sup>151</sup>

---

142 *Id.* at 1975.

143 *Id.*

144 *Id.*

145 *Id.* at 1976.

146 *Id.* at 1977-1980 (quoting *Conley v. Gibson*, 355 U.S. at 45-46).

147 *Id.* at 1978.

148 *Id.* at 1983.

149 *Id.*

150 *Id.*

151 *Id.* at 1988-1989.

## VI. THE FUTURE

*Twombly* is the latest in a line of recent Supreme Court cases that have not been kind to antitrust plaintiffs.<sup>152</sup> Indeed, the antitrust opinions from the past two terms show that this court will impose strict burdens of pleading and proof on plaintiffs in antitrust cases.<sup>153</sup> The court is heightening the burdens on plaintiffs by trending away from *per se* rules and toward analyses of the economic impacts of allegedly anticompetitive conduct. The court's reversal of the Second Circuit in *Twombly* is consistent with its reduced emphasis on blanket prohibitions of certain categories of conduct in favor of a more detailed inquiry into the nature and effect of the questioned conduct.

While *Twombly* is helpful to the extent it instructs plaintiffs on what not to plead, *i.e.*, parallel conduct with no facts to support an inference of an agreement, *Twombly* fails to provide meaningful guidance on what sorts of facts a plaintiff should plead to survive dismissal. The “plus factors” identified by Judge Posner at the summary judgment stage in *High Fructose Corn Syrup* may provide guidance to a plaintiff seeking to plead a conspiracy under Section 1.<sup>154</sup> In *High Fructose Corn Syrup*, Judge Posner noted several facts that the plaintiffs demonstrated to show the existence of “an explicit agreement, not merely a tacit one.”<sup>155</sup> Market structure allegations – few competitors, opportunities to communicate, ease of identification of cheating by other industry members – were certainly part of what made a finding of conspiracy factually permissible. However, there were also numerous express oral and written statements implying the existence of an illegal agreement, including direct statements from defendants’ management that the defendants had “an understanding . . . not to undercut each other’s prices,” and “our competitors are our friends.”<sup>156</sup> Most significantly, two former executives of the primary defendant (who were in prison for fixing the price of a different product), took the Fifth Amendment and refused to answer questions relating to whether they had fixed prices of high fructose corn syrup, permitting an adverse inference from their silence. The Seventh Circuit held that these facts represented “sufficient admissible evidence in support of the hypothesis of a price-fixing conspiracy to prevent the grant of summary judgment to the defendants.”<sup>157</sup>

As a practical matter, however, verbatim statements from defendants’ executives implying the existence of an agreement are neither commonly known – at least at the pleading stage – nor required to state a conspiracy claim under Section 1.<sup>158</sup> Perhaps the most significant plus factors “are those that tend to show that the conduct would be in the parties’ self-interest if all agreed to act in the same way but would be contrary to their self-interest if they acted alone.”<sup>159</sup> Allegations of this sort were missing from the Amended Complaint in *Twombly*, and, as noted by the district court and the Supreme Court, all of the plaintiffs’ conspiracy allegations were consistent with each defendant’s individual economic interest.<sup>160</sup>

The outcome of *Twombly* is broadly consistent with the views advanced by Professor Turner in his seminal article, in which he states:

The point is that conscious parallelism is never meaningful by itself, but always assumes whatever significance it might have from additional facts. Thus,

152 *E.g., Verizon Communications, Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004) (rejecting monopolization claim under Sherman Act Section 2, 15 U.S.C. Section 2, against Verizon based on Verizon’s purported failure to comply with the Telecommunications Act of 1996). Because the named plaintiffs in *Twombly* (and their lawyers, who represented the *Trinko* plaintiffs) were attempting to sidestep *Trinko*, one explanation for the Supreme Court’s disposal of the *Twombly* Amended Complaint would be the court’s desire to finish the job it had begun in *Trinko*. See Petition for a Writ of Certiorari, *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 550 U.S. \_\_\_\_ (No. 05-1126), 2006 WL 558418 (March 6, 2006), at \*4.

153 See *Volvo Trucks North America, Inc. v. Reeder-Simco GMC, Inc.*, 126 S. Ct. 860 (2006) (price discrimination); *Texaco, Inc. v. Dagher* 126 S. Ct. 1276 (2006) (joint ventures); *Illinois Tool Works Inc. v. Independent Ink, Inc.*, 126 S. Ct. 1462 (2006) (tying); *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc.*, 127 S. Ct. 1069 (2007) (predatory bidding); *Leegin Creative Leather Prods., Inc. v. PSKS, Inc.*, 127 S. Ct. 2705 (2007) (resale price maintenance).

154 *In Re High Fructose Corn Syrup Antitrust Litig.*, 295 F.3d 651 (7th Cir. 2002) (Posner, J.).

155 *Id.* at 661.

156 *Id.* at 662-63.

157 *Id.* at 666.

158 See, e.g., *In re: Pressure Sensitive Labelstock Antitrust Litig.*, 356 F. Supp. 2d 484, 492 (M.D. Penn. 2005).

159 American Bar Ass’n Section of Antitrust Law, *Antitrust Law Developments* 12 (6th ed. 2007).

160 *But see In re: Pressure Sensitive Labelstock Antitrust Litig.*, 356 F. Supp. 2d at 492 (declining to dismiss conspiracy complaint in which the behavior that was allegedly inconsistent with each defendant’s individual economic interest consisted of: 1) the existence of excess capacity in the labelstock industry; 2) an alleged lack of price competition; and 3) a statement by one defendant’s CEO that price competition from a new entrant had “ruined the market,” citing *Lum and Boggsian*). Faegre & Benson LLP represented two of the *Labelstock* defendants in a related federal merger enforcement action.

conscious parallelism is not even evidence of agreement unless there are some other facts indicating that the decisions of the alleged conspirators were interdependent, that the decisions were consistent with the individual self-interest of those concerned only if they all decided the same way.<sup>161</sup>

*Twombly's* result is also consistent with Judge Posner's recognition that there are practical limits to expanding the reach of Section 1 to encompass litigating complaints that allege no more than conscious parallelism:

In closing, might I suggest that if there is a good case to be made against the proposed employment of section 1 of the Sherman Act, it will be built on the practical difficulties of proving tacit collusion. That, rather than the doctrinal or remedial questions, is the heart of the problem and should be the focus of the debate.<sup>162</sup>

Thus, in *Twombly*, the Supreme Court premises its conclusion heavily upon fear of the overdeterrence that could result from allowing complaints pleading little more than similar actions by similarly situated competitors to survive motion to dismiss, when (in light of the court's refusal to reconsider *Theatre Enterprises*) such cases are substantively doomed.

The district court, the Second Circuit, and the Supreme Court each highlighted the problematic and burdensome discovery inherent in antitrust litigation. As noted, the Second Circuit went so far as to expressly invite the Supreme Court to "re-calibrate" the balance between the "colossal expense" of antitrust discovery and the ripple effect of the settlement of meritless claims by defendants to avoid this expense.<sup>163</sup>

In *Twombly*, the Supreme Court was surprisingly accepting of the Second Circuit's invitation to relieve the discovery burdens on antitrust defendants, as evinced by the court's holding and its devotion of several pages in its opinion to the burdensome nature of discovery and the case management challenges common to antitrust litigation. Standing alone *Twombly* need not mark a paradigm shift in discovery practice because it could be limited to a fairly narrow set of cases – Section 1 conspiracy claims which purport to premise an illegal agreement solely on parallel conduct among competitors. However, such limited application seems unlikely. After *Twombly* has had a few years to percolate downward through the federal court system, it may be identified as beginning an era of new limits on discovery in large civil cases, just as *Matsushita* ushered in a new era of judicial favor for summary judgments.

*Twombly's* greatest impact will occur as lower courts apply the plausibility standard under Rule 12 of the Federal Rules of Civil Procedure to dismiss cases beyond the realm of conspiracy claims founded on parallel conduct. If, as the majority insists, *Twombly* is not setting a heightened pleading standard unique to antitrust claims, then logically the case's reasoning applies to civil litigation in general. By reaching out to disparage *Conley v. Gibson*, the Supreme Court must be seen as allowing significantly freer rein to lower courts to dismiss complaints at the outset than has existed in recent memory.

161 Turner, *supra* note 1, at 658.

162 Posner, *supra* note 1, at 1592.

163 Discovery is nonetheless an integral part of litigation, and courts are generally reluctant to provide relief to parties complaining of discovery burdens outside of the protections afforded by the Federal Rules of Civil Procedure. The recent amendments to the Federal Rules governing electronic discovery arguably increase the scope and expense of discovery for corporate defendants. Even though these amendments excuse the non-production of materials because of "undue burden or cost," courts have generally refused to afford corporate defendants blanket relief from the burdens of discovery. Factually, when the *Twombly* defendants assert that they have collectively spent "billions on regulatory compliance efforts, and having their market-opening efforts exhaustively scrutinized by the Department of Justice and by federal and state regulatory authorities," the burden of one more civil lawsuit does not appear so great. See Petition for a Writ of Certiorari, *Bell Atlantic Corp. v. Twombly*, 127 S. Ct. 1955, 550 U.S. \_\_\_\_ (No. 05-1126), 2006 WL 558418 (March 6, 2006), at \*4.

# THE SEDONA CONFERENCE®

## INTRODUCTION TO THE LEGAL AND ECONOMIC ISSUES AT THE INTERSECTION OF THE PATENT AND ANTITRUST LAWS

---

*August 2007 Public Comment Version*

Editor:  
Thomas Greene\*

Patents represent strategic assets in the 21st century. But despite the importance of patents—or perhaps because of it—patent policy is increasingly unsettled. On the one hand, legal protection of new and novel ideas is the lifeblood of a modern economy. On the other, the modular nature of much innovation means that an old patent can hamper or block development of the next generation of technology. And overlapping thickets of patents must be navigated to make virtually any complex product. Ultimately, too much protection risks future advances while too little protection jeopardizes today's innovations.

Finding a path of grace between these two extremes is not a new challenge. Judges and legislatures have been wrestling with the appropriate mix of robust competition and state-sanctioned monopoly to advance new technology since at least the 16th century. In today's courtrooms, the boundaries of these two very different legal regimes are most often delineated at the points of intersection between patent law and antitrust law.

Working Group 4 of The Sedona Conference has been studying the intersection of patent and antitrust law. This article sketches the background for the Working Group's evaluation of this intersection. It reviews the common law roots of both antitrust and patent law, and highlights the potential for tensions developing between antitrust and intellectual property law.

### 1. Common Law Roots

The English common law was skeptical, if not hostile, to any form of monopoly. Sir Thomas More, for example, in his book *Utopia*, written in 1516, opined that “Suffer not thies ryche men to bye up all, to ingross and forstalle, and with theyr monopolye to kepe the market alone as please them.”<sup>1</sup> The royal family, however, could issue letters patent<sup>2</sup> to give individuals monopolies over particular lines of commerce.<sup>3</sup> This was relatively rare until the time of Elizabeth I who issued letters patent on a wide range of ordinary consumer goods including salt, iron, playing cards, beer and various kinds of cloth.<sup>4</sup> In response to the “odium which arose from abuse” of royal grants,<sup>5</sup> Parliament enacted the Statute of Monopolies in 1623.<sup>6</sup> The Act prohibited “all monopolies” with one exception. That exception was for letters patent for a period not to exceed fourteen years for the “sole working or making of any new manufactures within the realm” to be granted to the “true and first inventor” but only if “not contrary to the law nor mischievous to the state by raising prices of commodities at home, or hurt of trade or generally inconvenient.”<sup>7</sup>

\* This paper is a project of The Sedona® Conference Working Group on the Intersection of the Patent and Antitrust Laws (WG4). Special thanks go to all of the WG4 members and observers who have contributed to the development of this piece. This document is for educational purposes only and is not a substitute for legal advice. The opinions expressed herein are consensus views of the Working Group and do not necessarily represent the views of any individual participants or authors or any of the organizations to which they belong or clients they represent

1 H. Fox, *Monopolies and Patents* 24 (1947)

2 Letters patent were public documents as opposed to letters close which were sealed.

3 See *Darcy v. Allein*, 11 Co. Rep. 84b, 77 Eng. Rep. 1260 (*K.B. 1602*) (Darcy was allowed to monopolize the sale of playing cards pursuant to a royal grant even though such monopolies were contrary to the common law.)

4 Miller, *The Case of the Monopolies—Some of Its Results and Suggestions*, 6 Mich. L. Rev. 1, 2 (1907)

5 4 W. Holdsworth, *A History of English Law* 348 (2nd ed. 1937)

6 An Act concerning Monopolies and Dispensations with Penal Laws and the Forfeitures thereof, 21 Jac. 1, c. 3 (4 Statutes at Large 734 (1811))

7 *Id.* at section 6

Sir Edward Coke's commentaries on the common law of England became the standard reference works for law students like Jefferson and Adams in the American colonies. His report on the Case of Monopolies concluded that a grant of an exclusive license to sell playing cards "was utterly void" as "against the common law."<sup>8</sup> He noted that among the "inseparable incidents" of every monopoly were that (1) "the price of the same commodity will be raised, for he who has the sole selling of any commodity, may and will make the price as he pleases"; (2) "after the commodity [is] granted, the commodity is not so good as it was before"; and (3) monopoly "tends to the impoverishment of divers artificers" who are precluded from making the monopolized product.<sup>9</sup>

When the U.S. Constitution was drafted, four states suggested amendments that would have paralleled the English Statute of Monopolies. Although advocated by Jefferson, this proposal was not adopted.<sup>10</sup> Rather, the new Constitution authorized Congress to enact laws to "promote science and the useful arts by securing for limited times to authors and inventors the exclusive right to their respective writing and discoveries."<sup>11</sup> According to Madison in Federalist No. 43, this was included to "protect a right of common law."<sup>12</sup> And in this instance, he argued, the "public good fully coincides...with the claims of individuals."<sup>13</sup>

The first patent law under the new Constitution was enacted in 1793 as "An Act to Promote the Progress of Useful Arts". This was succeeded in 1793 with a statute substantially attributed to Jefferson,<sup>14</sup> the language of which is virtually identical to parts of today's law.<sup>15</sup> Patentable subjects included "any new and useful art, machine, manufacture or composition of matter or any new and useful improvement on any art, machine, manufacture or composition of matter."<sup>16</sup> An applicant had to provide a "written description of his invention, and of the manner of using, or process of compounding the same, in such full, clear and exact terms to distinguish the same from all other things before known and to enable any person skilled in the art or science of which it is a part, or with which it is most nearly connected, to make, compound and use the same."<sup>17</sup> An inventor was required to "fully explain the principle, and the several modes in which he has contemplated the application of" his invention.<sup>18</sup> Only the "true inventor" could seek to obtain a patent and only if the invention had not been "known or used before the application."<sup>19</sup> Like the 1623 Statute on Monopolies, letters patent guaranteed a state-protected monopoly for fourteen years.

The nineteenth century saw a dramatic increase in the number of patents. Doctrinally, the most important legal development was the mid-century decision in *Hotchkiss v. Greenwood*,<sup>20</sup> that clarified that a patentable invention had to be not only new, but not obvious. This was enforced by a new examination system that is credited by economic historians with reducing the number of patent lawsuits and spurring innovation.<sup>21</sup> In 1887, the United States joined the Paris Convention for the Protection of Industrial Property to become a formal part of international efforts to protect patented inventions.

In 1890, the Sherman Act became law.<sup>22</sup> The new statute prohibited both combinations in restraint of trade and monopolization. Senator Sherman stated that the new Act was designed to "supplement the enforcement of established rules of the common and statute law by the courts of the several States."<sup>23</sup> With enactment of the Sherman Act, the competing elements of the Statute on Monopolies became part of U.S. statutory law. The general prohibitions against restraints of trade and monopoly were now explicit in the Sherman Act while its limited exception to encourage new inventions was captured in the Patent Act.

8 *Darcy v. Allein* (The Case of Monopolies) 11 Co. Rep. 84b, 77 Eng.Rep. 1260 (K.B. 1603)

9 *Id.*

10 Letter, Jefferson to Madison, (Aug. 28, 1789, in *The Republic of Letters* 1 (James Morton Smith, ed., 1995))

11 U.S. Constitution, Art. 1, Section 8, clause 8

12 The Federalist No. 43 (Madison)

13 *Id.*

14 *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 147 (1989)

15 Compare Patent Act of 1793, 1 Stat. 318-323 (February 21, 1793), section 1 with 35 U.S.C. section 101

16 Patent Act of 1793, section 1, 1 Stat. 318-323 (February 21, 1793)

17 Patent Act of 1793, section 3, 1 Stat. 318-323 (February 21, 1793)

18 *Id.*

19 Patent Act of 1793, sections 3 and 1, 1 Stat. 318-323 (February 21, 1793)

20 52 U.S. 248 (1850)

21 Zorina B. Khan, *The Democratization of Invention: Patents and Copyrights in American Economic Development, 1790-1920* (2004)

22 26 Stats. 209 (July 2, 1890)

23 21 Cong. Rec. 2547 (1890)

## 2. The PTO and its “Customers”

The Patent and Trademark Office issued 165,485 patents in federal fiscal year 2005. This is over twice the number issued in 1985 and approximately 30% more than in 1995.<sup>24</sup> Despite political rhetoric suggesting that patents uniquely advance U.S. interests, 80,247 of these patents, or approximately half, were issued to residents of foreign nations.<sup>25</sup> Japan led with 34,079 patents with Germany second with 10,502. Emerging technology centers like China and India were well behind with 583 and 405 patents, respectively.

The PTO has declared its mission to be “helping our customers get patents.”<sup>26</sup> Examiners spend an estimated 18 hours on an individual application, reviewing the application, testing it against prior art contained in various data bases and writing up their analyses.<sup>27</sup> The process is secret and much of the burden of providing prior art, particularly in the most innovative industries, is on patentees’ legal representatives.<sup>28</sup> Incentives in the PTO are generally understood to favor grants over denials,<sup>29</sup> and 95-97% of all applications ultimately result in a patent.<sup>30</sup> Despite this approval rate, it takes an average of thirty-one months to process a patent application, a backlog the agency expects to close by hiring 1,200 new examiners in 2007.<sup>31</sup> These hires are in addition to a record-setting addition of 1,193 examiners in 2006.<sup>32</sup>

In the last 15 years, the PTO has administered a system that has become increasingly patent friendly. The United States Supreme Court concluded in 1980 that a newly developed microorganism could be the subject of a patent in *Diamond v. Chakrabarty*.<sup>33</sup> In the following year, it determined in *Diamond v. Diehr* that a computer program could be patented.<sup>34</sup>

The doctrine of equivalency was given a broad sweep in the Court’s decisions in *Warner-Jenkinson Co. v. Hilton Davis Chemical Co.*<sup>35</sup> and *Festo v. Shoketsu Kinzoku Kogyo Kabushiki*, although the Court acknowledged that this made the scope of patents “less certain.”<sup>36</sup> These decisions contrast with earlier cases that opined that clear definition of claims was necessary to “guard against unreasonable advantages to the patentee and disadvantages to others arising from uncertainty as to their rights.”<sup>37</sup>

The Federal Circuit, created in 1982 to bring uniformity to patent jurisprudence, raised evidentiary standards for challenging patents in 1986.<sup>38</sup> The same court relaxed standards for evaluating whether an invention is “obvious” to practitioners skilled in the art.<sup>39</sup> It also softened the “best mode” requirement, substantially freeing applicants from having to specify the means by which their inventions will work in the real world.<sup>40</sup> According to an IBM attorney, this “invites the patenting of ideas that may have been visualized as desirable but have no foundation in terms of the research or development that may be required to enable their implementation.”<sup>41</sup>

24 U.S. Patent and Trade Office, *Performance and Accountability Report Fiscal Year 2005*, Table 6: Patents Issued, available at [http://www.uspto.gov/web/offices/com/annual/2005/060406\\_table6.html](http://www.uspto.gov/web/offices/com/annual/2005/060406_table6.html)

25 *Id.* Table 10: Patents Issued by United States to Residents of Foreign Countries (FY 2001-FY 2005) (Preliminary for FY 2005), available at [http://www.uspto.gov/web/offices/com/annual/2005/060410\\_table10.html](http://www.uspto.gov/web/offices/com/annual/2005/060410_table10.html)

26 Mark Lemley, *Rational Ignorance at the Patent Office*, 95 Northwestern L. Rev. 1, 2, n. 3 (2001). The current formal mission statement provides that: “The USPTO’s mission is to insure that the Intellectual Property system contributes to a strong global economy, encourages investment in innovation, and fosters entrepreneurial spirit. Intellectual property is an invention or creation embodied in the form of a patent, trademark, trade secret, or copyright.” available at <http://www.uspto.gov/web/menu/intro.html>.

27 Brenda Sandburg, *Speed Over Substance? Intell. Prop. Mag.*, Mar 1999 (estimating 18 hours on average with more time spent on more complex applications); compare *Patent Nonsense: The Knowledge Monopolies*, *The Economist*, April 8, 2000 (“[P]atent examiners spend only eight hours on a patent, on average.”)

28 37 CFR section 1.56

29 Mark Lemley, *Rational Ignorance at the Patent Office*, 95 Northwestern L. Rev. 1, 2, n. 3 (2001)

30 Cecil D. Quillen, Jr. & Ogden H. Webster, *Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office*, 11 Fed. Cir. B.J. 1 (2001)

31 U.S. Patent and Trademark Office, 2007-2012 Strategic Plan 6, available at <http://www.uspto.gov/web/offices/com/strat2007/stratplan2007-2012.pdf>

32 *Id.*

33 447 U.S. 303 (1980)

34 450 U.S. 175 (1981)

35 520 U.S. 17 (1997)

36 535 U.S. 722, 732 (2002)

37 *McClain v. Ortmyer*, 141 U.S. 419, 424 (1891), see also *General Electric Co. v. Wabash Appliance Corp.*, 304 U.S. 364, 369 (1938)

38 *Medtronics, Inc. v. Intermedics, Inc.*, 799 F.2d 734 (Fed. Cir. 1986); *Hybridtech Inc. v. Monoclonal Antibodies Inc.*, 802 F.2d 1367 (Fed. Cir. 1986)

39 *Sratoflex, Inc. v. Aeroquip Corp.*, 713 F.2d 1530 (Fed. Cir. 1983); *Simmons Fastener Corporation v. Illinois Tool Works*, 739 F.2d 1573 (1984)

40 Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology Specific?*, 17 Berkeley Technology Law Journal 9-10 (2003)

41 John D. Flynn, *Comments on the International Effort to Harmonize the Substantive Requirements of Patent Laws* (IBM 2001), available at <http://www.uspto.gov/web/offices/dcom/olia/harmonization/TAB42.pdf>

Finally, the Federal Circuit decided in *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*<sup>42</sup> that “business methods” are patentable. As one patent specialist noted in the *National Law Journal*, companies should “now seek U.S. patent rights for any conceivable business operation, such as methods of billing clients, hiring employees, marketing products or service...or simply obtaining funding.”<sup>43</sup>

These developments have engendered dramatically different reactions. At one end of the spectrum, the PTO argues that the current system “has propelled our nation from a small agrarian society to the preeminent technological and economic superpower...and has become the basis for economic development in nations throughout the world.”<sup>44</sup> At the other, the National Institutes of Health have stated categorically that the granting of patent rights for biological research tools, expected to be critical for the development of stem cell technologies, “can stifle the broad dissemination of new discoveries and limit future avenues of research and product development.”<sup>45</sup>

In the middle, a 2003 report of the Federal Trade Commission concluded that while there was “much to praise” in the system, “[p]oor patent quality and legal standards and procedures that inadvertently may have anticompetitive effects can...hamper competition that otherwise would stimulate innovation.”<sup>46</sup> To address its concerns, the FTC suggested a number of reforms including third-party involvement in challenging patents during the examination process, strengthening the obviousness requirement and reclaiming the “preponderance of the evidence” review standard.<sup>47</sup>

Likewise, the National Academy of Sciences concluded in a major report in 2004 that “[c]ontinuing high rates of innovation suggest that the patent system is working well and does not require fundamental changes.”<sup>48</sup> However, the National Academy, like the FTC, suggested that a number of reforms are necessary including reinvigoration of the obviousness requirement, institution of third-party participation in the process at the PTO, shielding some research uses of patented products or processes from infringement claims and increased staffing at the agency.<sup>49</sup>

Both the Supreme Court and the Congress have begun to react to concerns about patent quality. In *KSR International v. Teleflex, Inc.*,<sup>50</sup> the Court rejected Federal Circuit precedent on obviousness, scoring that court’s approach as too formulaic. The Supreme Court concluded that:

...the results of ordinary innovation [based on existing art] are not the subject of exclusive rights under the patent laws. Were it otherwise patents might stifle, rather than promote, the progress of useful arts.<sup>51</sup>

Other high court cases have also begun to cut back on Federal Circuit decisions. In two important procedural decisions, the Court has made it easier to effectively challenge poor patents. In *eBay, Inc. v. MercExchange, L.L.C.*,<sup>52</sup> the Court rejected the principle that injunctions should be issued in the normal course of patent litigation. It noted that such a rule was contrary to equity jurisprudence in other areas of the law, and four justices argued that the reflexive issuance of injunctions gave “undue leverage” to patent holders for claims of “potential vagueness and suspect validity.”<sup>53</sup> In *Medimmune, Inc. v. Genentech, Inc.*,<sup>54</sup> the Court rejected jurisprudence requiring a patent licensee to breach its licensing agreement in order to seek a judgment on the patent’s validity.

42 927 F. Supp. 502, 516 (D.Mass. 1996), rev’d, 149 F.3d 1368 (Fed. Cir. 1998), cert. denied, 525 U.S. 1093 (1999)

43 Barry Schindler, *In Focus: Intellectual Property. Key ruling for business methods*, *The National Law Journal* (December 5, 2005)

44 Jon W. Dudas, Message from the Undersecretary of Commerce for Intellectual Property and Director of the U.S. Patent and Trademark Office 1 (November 2, 2005, available at [http://www.uspto.gov/web/offices/com/annual/2005/02\\_message\\_director.html](http://www.uspto.gov/web/offices/com/annual/2005/02_message_director.html))

45 Principles and Guidelines for Recipients of NIH Grants for Contracts on Obtaining and Disseminating Biomedical Research Resources: Final Notice, 64 Fed.Reg. 72,090, 72,092 (December 23, 1999); see also Scott Iyama, *The USPTO’s Proposal of a Biological Research Tool Patent Pool Doesn’t Hold Water*, 57 Stan.L.Rev. 1223 (2005)

46 Federal Trade Commission, *To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy* 4-5 (October 2003), available at <http://www.ftc.gov/os/2003/10/innovationrpt.pdf>

47 *Id.* at 7-18

48 Committee on Intellectual Property Rights, National Research Council, *A Patent System for the 21st Century* 1 (Stephen A. Merrill, Richard C. Levin & Mark B. Myers, eds.) (National Academy of Sciences, 2004)

49 *Id.* at 5-8.

50 \_\_\_ U.S. \_\_\_, 127 S.Ct. 1727 (2007)

51 *Id.* at 1746.

52 \_\_\_ U.S. \_\_\_, 126 S.Ct. 1837 (2006).

53 *Id.* at 1842.

54 549 U.S. \_\_\_, 127 S.Ct. 764 (2007).

Congress has begun its own overhaul of the patent system. After several years of hearings, omnibus bills to amend the patent law have been reported to the floors of both houses.<sup>55</sup> Although there are major differences in the two bills, both the House and Senate bills provide for a post-grant opposition procedure already used in Europe to improve patent quality.<sup>56</sup>

### 3. Patent Thickets, Strategic Portfolios and the Business of Innovation

Many people think that a single patent will give rise to a complete product, as Alexander Graham Bell's work gave rise to the telephone. Modern realities are far different. The DVD player and disk, for example, require the interplay of 115 patents for the players and 95 patents for the disks themselves. The patents for the practical production of these products are held by Koninklijke Philips Electronics, N.V. Sony Corporation of Japan and Pioneer Electronics of Japan.<sup>57</sup>

The interplay of patents can be so dense that innovators can face a "patent thicket". Such thickets have been defined academically as "an overlapping set of patent rights requiring that those seeking to commercialize new technology obtain licenses from multiple patentees."<sup>58</sup> Patent thickets for innovative companies are akin to walking through a dense wood without stepping on a twig.

In response to patent thickets in which multiple patents from competing companies "read on" each other, technology companies have sought collections of patents that will force other companies in the thicket to seek licenses from them. The key is the ability to threaten others who wish to make products with potential infringement claims. According to Cisco general counsel Mark Chandler, his company invests in patents "to assure that if someone wants to assert patents against us, we will have some countervailing tools."<sup>59</sup> Research suggests that large companies tend to seek a large number of patents while smaller companies will seek fewer, but more strategic, patents.<sup>60</sup>

The value of a patent portfolio is hard to measure. However, an effective portfolio should facilitate in-house innovation, side-step expensive litigation, improve the company's bargaining position with rivals and enhance efforts to attract capital.<sup>61</sup> A closely related set of patents can operate as a "super-patent", fencing competitors out of a lucrative area.<sup>62</sup>

The cost of securing a U.S. patent in 1996 ranged from \$10,000 to \$30,000.<sup>63</sup> The cost of securing the same patent in 10 European countries was typically \$95,000.<sup>64</sup> If the patents are litigated, a survey of the American Intellectual Property Law Association in 2003 found that the median cost of discovery in actions involving less than \$1 million was \$290,000 while the total litigation costs were \$500,000.<sup>65</sup> For actions in which the patent was worth between \$1 million and \$25 million, discovery costs were \$1 million and the total litigation costs were \$2 million.<sup>66</sup> While these costs are substantial, and have certainly risen, even a spurious threat of an infringement claim can cause companies to pay significant sums. For example, hundreds of companies paid a total of \$1.5 billion in royalties to the Lemelson Foundation for so-called "submarine" patents on bar code technology that were ultimately held to be unenforceable by the Federal Circuit.<sup>67</sup>

55 H.R. 1908 (Berman et al.), reported to the House floor on July 18, 2007, S. 1145 (Leahy et al.), reported to the Senate floor on July 19, 2007; See Marcia Coyle, *Patent reform finds traction*, 29 National Law Journal 1 (July 30, 2007)

56 *Id.* at 17

57 DOJ Business Review Letter, Assistant Attorney General Joel I. Klein to Gerard R. Beeny, re: Proposed package licensing of essential DVD patents 2 (December 16, 1998), available at <http://www.usdoj.gov/atrp/public/busreview/2121.htm>

58 Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools and Standard-Setting*, 1 *Innovation Policy and the Economy* 1 (Adam Jaffe, Joshua Lerner and Scott Stern, eds.) (MIT Press, 2001)

59 Michael Orley, *The Patent Epidemic: It's wasting companies' money and slowing the development of new products*, Business Week (January 6, 2006) 60, 61

60 John R. Allison & Mark A. Lemley, *Who's Patenting What? An Empirical Exploration of Patent Prosecution*, 53 *Vanderbilt L.Rev.* 2099, 2128 (2000); see also FM Scherer, *Schumpeter and Plausible Capitalism*, 30 *J.Econ. Literature* 1416, 1423 (1992)

61 Gideon Parchomovsky & R.Polk Wagner, *Patent Portfolios*, 154 *U.Pa.L. Rev.* 1, 33-44 (2005); see also James Bessen, *Patent Thickets: Strategic Patenting of Complex Technologies*, available at <http://www.researchoninnovation.org/online.htm#thicket>

62 *Id.* 32-33

63 Wayne M. Kennard, *Obtaining and Litigating Software Patents*, 431 *PLI/Pat* 193, 208 (1996)

64 Edwin F. Berrier, Jr., *Global Patent Costs Must Be Reduced*, 36 *IDEA* 473, 476-77 (1996)

65 Am. Intellectual Prop. Law Assoc., *Report of the Economic Survey 2003*, at 22 (2003), cited in James E. Besson & Michael J. Meurer, *Lessons for Patent Policy from Empirical Research on Patent Litigation*, 9 *Lewis & Clark L.Rev.* 1, 2, n. 5

66 *Id.*

67 Brenda Sandburg, *Lemelson patents are unenforceable*, *The Recorder* (September 13, 2005); see *Symbol Technologies v. Lemelson Medical, Education & Research Foundation*, 422 F.3d 1378 (Fed. Cir. 2005)



#### 4. Academic Perspectives on Patents and Innovation

The role of innovation in the economy is an on-going subject of academic research and thought. Joseph Schumpeter, an early writer on innovation, famously rejected much classical economic thought, arguing that “perennial gales of creative destruction” made concepts like market power of limited relevance.<sup>68</sup> At the other end of the spectrum, it has been argued that patent monopolies—assuming they confer market power—must be closely circumscribed.<sup>69</sup> This view was reflected in the so-called “nine no-no’s”, a list of licensing practices that the U.S. Department of Justice once considered presumptively unlawful.<sup>70</sup>

By the 1970’s, it was argued that since patents can add to consumer welfare, there was no necessary tension between antitrust and patent law.<sup>71</sup> Ward Bowman, for example, wrote that:

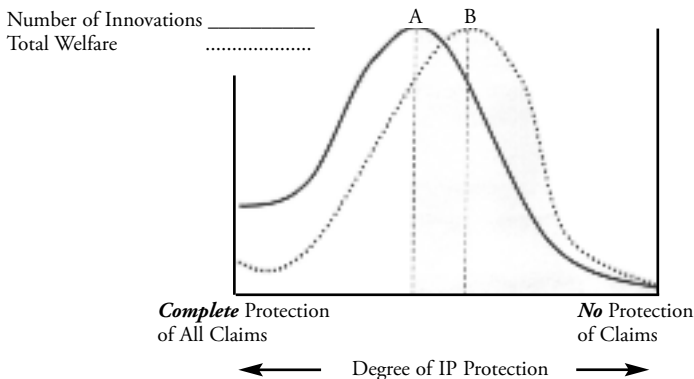
Both antitrust law and patent law have a common central economic goal: to maximize wealth by producing what consumers want at the lowest cost. In serving this common goal, reconciliation between patent and antitrust law involves serious problems of assessing effects, but not conflicting purposes.<sup>72</sup>

While Bowman saw this as theoretically correct, he cautioned that there are “serious problems of assessing effects” of patent monopolies.

Landes and Posner commented that “if intellectual property rights are enforced too strictly, then subsequent innovators will be foreclosed and overall welfare will be reduced.”<sup>73</sup> Likewise, a respected Berkeley economist has noted: “when there are multiple gate keepers, each of whom must grant permission before a resource can be used...the resource may be underutilized” and, in the case of patents, “innovation is stifled.”<sup>74 75</sup>

The point made by all three of these commentators is illustrated by a chart created by James Langenfeld.<sup>76</sup> In Figure 1, both the number of innovations and total welfare are charted on a single graph. The horizontal axis delineates the possible levels of IP protection from complete IP protection at the far left to no protection at the far right.

**Figure 1: Optimal IP Protection**



68 Joseph A. Schumpeter, *The Process of Creative Destruction* (Unwin, 1942).

69 See, e.g., *Mercoid Corp. v. Mid-Continent Inv. Co.*, 320 U.S. 661 (1944); *Carbice Corp. v. American Patents Dev. Corp.*, 283 U.S. 27 (1931)

70 For a description of the Nine No-No’s, see Willard K. Tom & Joshua A. Newberg, *Antitrust and Intellectual Property: From Separate Spheres to Unified Field*, 66 *Antitrust L. J.* 167, 178-184 (1998).

71 Charles Rule, *The Administration’s Views: Antitrust Analysis after the Nine No-No’s*, 55 *Antitrust L.J.* 365 (1986); see also Richard Gilbert and Carl Shapiro, *Antitrust Issues in the Licensing of Intellectual Property: The Nine No-No’s Meet the Nineties*, Brookings Papers on Economics: Microeconomics 283-336 (1997)

72 Ward Bowman, Jr., *Patent and Antitrust Law: A Legal and Economic Portfolio* (1973)

73 William Landes & Richard Posner, *An Economic Analysis of Copyright Law*, 18 *J. Legal Stud.* 325, 326 (1989)

74 Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools and Standard-Setting*, 1 *Innovation Policy and the Economy* (Adam Jaffe, Joshua Lerner, and Scott Stern, eds. MIT Press, 2001)

75 Researchers have also uncovered a so-called “patent paradox”, that is, increased patenting is associated with declining expenditures on research and development and reductions in real innovation. Gideon Parchomovsky & R. Polk Wagner, *Patent Portfolios*, 154 *U.Pa.L. Rev.* 1 (2005)

76 James Langenfeld, *Intellectual Property Protection and Antitrust: Steps Toward Striking a Balance*, 53 *Case Western Res. L.Rev.* 91, 97 (2001)

Looking first at innovation, total innovation is lower with complete protection of all IP claims because subsequent inventors are foreclosed from prior art unless they pay significant rents. This is reflected by the point of intersection between the total innovation line and the vertical axis. But innovation increases to point A as innovators are allowed to more fully take advantage of prior art. After Point A, however, innovation declines because the inventors' incentives are diminished. That is, inventors cannot reap as many profits from their inventions because others can more easily copy the innovation and drive down prices.

The line for total welfare peaks at Point B, somewhat to the right of the peak of the total innovation line. According to Dr. Langenfeld:

With complete intellectual property protection, total welfare is relatively lower than the number of innovations. Innovators gain all of the benefits from their innovations, there is no price competition or competition from the follow innovations of others, and no consumer surplus from innovations. As intellectual property protection is relaxed (moving left to right in Figure 1), total welfare increases to its peak at point B, with more development innovations by others and more competition reducing prices and increasing consumer welfare. The optimal total welfare will in general be at the point B, right of point A, indicating that total welfare is maximized with less intellectual property protection than a structure designed to maximize innovations. However, reducing intellectual property protection below point B reduces total welfare as innovators have increasingly less incentive to innovate and fewer innovations occur.<sup>77</sup>

Antitrust enforcers have typically focused on maximizing total or consumer welfare,<sup>78</sup> not maximizing the number of innovations. If patent law is interpreted to maximize the number of innovations and not weigh the benefits of price competition to consumers, then there is the clear possibility that patent and antitrust laws can come into conflict. This is an important touchstone when assessing the competing, and often contentious, claims of patent and antitrust law.

This analysis assumes that there is a clear relationship between the scope and depth of intellectual property protection and the production of new innovations. This assumed relationship is the subject of increasing scrutiny, with some economists arguing that patents play a "surprisingly minor role" in innovation pointing, instead, to larger economic trends like levels of education and the scope of public funding of basic research.<sup>79</sup> Recognizably, such research could have a profound effect on how we think about the relationship between antitrust law and the patent system.

## 5. Conclusion

In the context of the potential tension of patent and antitrust laws, the Sedona Working Group 4 is in the process of analyzing a number of specific recurring and practical issues at the intersection of antitrust and patent laws. The Working Group is composed of lawyers and economists, members of plaintiff and defense bars, intellectual property and competitive experts, and public officials. This group is attempting to develop principles of decision and analysis that offer practical guides to navigating the particularly difficult points at the intersection between antitrust and patent law.

<sup>77</sup> Langenfeld, *supra*, n. 68 at 98. The precise peaks of the curves discussed above will likely be the subject of further empirical research. Economists continue to research the sources and drivers of innovation in modern economies. In particular, research continues on the relationship of innovation to levels of IP protection, levels of education, macroeconomic activity and many other factors.

<sup>78</sup> See, e.g., Timothy J. Muris, *Robert Pitofsky Public Servant and Scholar*, 52 CASE W. RES. L. REV. 25, 37 (2001). ("there is wide-spread agreement that the purpose of antitrust is to protect consumers"). Antitrust laws allow government agencies or private parties to obtain relief by eliminating practices that reduce competition in pricing or innovation, and obtaining fines or damages.

<sup>79</sup> This analysis assumes that the degree of IP protection is the important driver of innovation. Various writers suggest that many other factors affect innovation. See, e.g., F.M. Scherer, *The Political Economy of Patent Policy Reform in the United States* (Harvard University, 2006), available at <http://scholar.google.com/scholar?hl=en&lr=&oi=q&q=fin+scherer+the+political+economy+of+patent+policy+author:f+scherer> (patents play "surprisingly minor role" in decisions of companies to invest in research and development)



# CLAIM CONSTRUCTION AND IMPLICIT DEFINITIONS BASED ON THE SPECIFICATION SINCE *PHILLIPS*

---

*Robert Fram, Laurie Adams & Tanya Mazur  
Heller Ehrman  
San Francisco, CA*

## I. INTRODUCTION

Since the Federal Circuit's decision in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005), *cert. denied*, 126 S.Ct. 1332 (2006), the Court has continued to struggle with implementing a framework for claim construction. Although many had hoped that *Phillips* would bring clarity and consistency to the methodology of claim construction, that hope has not yet fully materialized. Nonetheless, certain trends have developed. In particular it appears that, since *Phillips*, the Federal Circuit is more likely than not to find that the specification supports an implicit definition of a claim limitation.<sup>1</sup>

To be sure, this line of thought resonates with greater or lesser force with different panels. Despite this, it is worthwhile to note some of the circumstances in which an argument for including an implicit definition have been particularly successful. In particular:

- (1) when a patent makes consistent use of a term throughout the specification;
- (2) when the preferred embodiment is referred to as “the invention;”
- (3) when a structure is disclosed in the specification that, by its nature, excludes others;
- (4) when the specification distinguishes the invention from the prior art on the basis of a particular feature;
- (5) when the specification demeans a technology or solution, and;
- (6) reliance on industry standards, when incorporated in the specification, as highly probative of the meaning of a term.

Section I of this paper briefly reviews the manner in which the Federal Circuit had addressed the role of the specification before *Phillips*. Section II reviews the framework articulated in *Phillips*. Section III reviews the post-*Phillips* jurisprudence on the question of implicit definition based on the specification, reviewing the six arguments set forth above. It concludes with a discussion of the persisting question of the “single embodiment” specification in the wake of *Phillips*.

## II. PRE-*PHILLIPS* CLAIM CONSTRUCTION JURISPRUDENCE

The use of the specification in construing the meaning of claims has always been bounded by two fundamental principles which are often in tension with each other: “(a) one may not read a

---

<sup>1</sup> To be sure, there are several other areas of inquiry that bear on claim construction (e.g., the appropriate role of extrinsic evidence) that continue to be the focus of the Federal Circuit's attention. That those subjects are beyond the scope of this paper does not diminish their importance.

limitation into a claim from the written description, but (b) one may look to the written description to define a term already in a claim limitation.” *Renishaw PLC v. Marposs Societa per Azioni*, 158 F.3d 1243, 1248 (Fed. Cir. 1998). In the years leading up to *Phillips*, two divergent approaches to claim construction arose, the dictionary-based methodology and the specification-based methodology.

The dictionary-centric approach was brought to the fore in *Texas Digital Systems, Inc. v. Telegenix, Inc.* 308 F.3d 1193 (Fed. Cir. 2002). The Court in *Texas Digital* stated that dictionaries and other related materials “may be the most meaningful sources of information to aid judges in better understanding both the technology and the terminology used by those skilled in the art to describe the technology.” *Id.* at 1203. The *Texas Digital* court recommended the use of the intrinsic record to determine which of multiple dictionary meanings “is most consistent with the use of the words by the inventor.” Indeed, the Court warned that “[c]onsulting the written description and prosecution history as a threshold step in the claim construction process, before any effort is made to discern the ordinary and customary meanings attributed to the words themselves, invites a violation of our precedent counseling against importing limitations into claims.” *Id.* at 1204. The presumption in favor of dictionary meanings would only be overcome “where the patentee, acting as his or her own lexicographer, has clearly set forth an explicit definition of the term different from its ordinary meaning.” *Id.*

The specification-centric approach instead focused on the specification and prosecution history to determine the proper meaning of disputed claim terms. Thus, in *Toro Co. v. White Consolidated Industries, Inc.*, the Federal Circuit took a very different approach:

[W]ords of ordinary usage must...be construed in the context of the patent documents. Thus the court must determine how a person of experience in the field of the invention would, upon reading the patent documents, understand the words used to define the invention.... As this case well illustrates, the dictionary definitions of common words are often less useful than the patent documents themselves in establishing the usage of ordinary words in connection with the claimed subject matter.

199 F.3d 1295, 1299 (Fed. Cir. 1999).

Torn between these divergent approaches, the Federal Circuit, lower courts, lawyers, and litigants struggled with how properly to construe disputed claim terms. It was against this background that the Federal Circuit agreed to consider *Phillips en banc*.

### III. *PHILLIPS V. AWH CORP.*

The *Phillips en banc* decision was widely seen as promising the definitive word by the Federal Circuit regarding claim construction. The well known facts of the case can be briefly stated. *Phillips* owned a patent for vandalism-resistant wall panels, and sued AWH for infringement. The United States District Court for the District of Colorado granted summary judgment for noninfringement based in part on its construction of the claim term “baffles.” Finding that “every textual reference in the Specification and its diagrams show baffle deployment at an angle other than 90 to the wall faces,” the District Court concluded that “for purposes of the . . . patent [at issue], a baffle must ‘extend inward from the steel shell walls at an oblique or acute angle to the wall face’ and must form part of an interlocking barrier in the interior of the wall module.” *Id.* at 1309 (citation omitted).

On appeal, a Federal Circuit panel affirmed, agreeing with the District Court that the patent used the term “baffles” in a restrictive manner. *Id.* at 1310. Based on the patent’s written description, the panel held that the claim term ‘baffles’ excludes structures that extend at a 90 degree angle from the walls. The panel also found that “the patent specification ‘is intended to support and inform the claims, and here it makes it unmistakably clear that the invention involves baffles angled at other than 90.’” *Id.* (quoting *Phillips v. AWH Corp.*, 363 F.3d 1207, 1213 (Fed. Cir. 2004)). Judge Dyk dissented, stating

that the panel “improperly limited the claims to the particular embodiment of the invention disclosed in the specification, rather than adopting the ‘plain meaning’ of the term ‘baffles.’” Instead, Judge Dyk wrote, the Court should have adopted a dictionary definition. *Id.*

Upon rehearing the case en banc, the Federal Circuit vacated the panel’s judgment and ruled, instead, that proper construction of the term “baffles did not preclude baffles disposed at a 90 degree angle.” Before providing its analysis of the proper construction of the term, the Federal Circuit said that it was “restating” and “reaffirming” the basic principles of claim construction outlined in *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 979-81 (Fed. Cir. 1995) (en banc), *aff’d*, 517 U.S. 370 (1996); *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576 (Fed. Cir. 1996); and *Innova/Pure Water, Inc. v. Safari Water Filtration Systems, Inc.* 381 F.3d 1111 (Fed. Cir. 2004). The Court also reiterated the importance of reading the claim terms “not only in the context of a particular claim in which the disputed term appears, but in the context of the entire patent, including the specification.” *Phillips*, 415 F.3d at 1313. The Court noted that in certain cases, “the ordinary meaning of claim language as understood by a person of skill in the art may be readily apparent even to lay judges . . . .” *Id.* at 1314. It is only when the meaning of a term is not readily understood by persons of skill in the art or the patentee has used a particular term idiosyncratically that a court should look to (1) “the words of the claims themselves,” (2) “the remainder of the specification,” (3) “the prosecution history,” and (4) “extrinsic evidence concerning relevant scientific principles, the meaning of technical terms, and the state of the art.” *Id.* (quoting *Innova/Pure Water, Inc. v. Safari Water Filtration Sys., Inc.*, 381 F.3d 1111, 1116 (Fed. Cir. 2004)).

The Court stressed that “[t]he best source for understanding a technical term is the specification from which it arose, informed, as needed, by the prosecution history.” *Id.* at 1315 (quoting *Multiform Desiccants, Inc. v. Medzam, Ltd.*, 133 F.3d 1473, 1478 (Fed. Cir. 1998)). Further, the Court explained, “the construction that stays true to the claim language and most naturally aligns with the patent’s description of the invention will be, in the end, the correct construction.” The Court criticized limiting the use of the specification and the requirement that “any definition of claim language in the specification be express” as “inconsistent with our rulings that the specification is ‘the single best guide to the meaning of a disputed term,’ and that the specification ‘acts as a dictionary when it expressly defines the terms used in the claims or when it defines terms by implication.’” *Id.* at 1320-21 (quoting *Vitronics*, 90 F.3d at 1582). The Federal Circuit was concerned that this line of reasoning would “extend patent protection beyond what should properly be afforded by the inventor’s patent” and would interfere with the public notice function.

Finally, the Court acknowledged that:

[t]he manner in which the patentee uses a term within the specification and claims usually will make the distinction” between whether “the patentee is setting out specific examples of the invention . . . or . . . the patentee instead intends for the claims and the embodiments in the specification to be strictly coextensive.

*Id.* at 1323.

*Phillips* presents a variety of analysis to support reliance on implicit definitions found in the specification. The most persuasive analysis found in *Phillips* to acknowledge the “revival” of implied definitions is found in the Court’s discussion of the *Texas Digital* case. While discussing that case, the Federal Circuit noted that “[a]ssigning such a limited role to the specification, and *in particular requiring that any definition of claim language in the specification be express*, is inconsistent with our ruling that the specification is ‘the single best guide to the meaning of a disputed term,’ and that the specification ‘acts as a dictionary when it expressly defines terms used in the claims or when it defines terms by implication.’” *Id.* at 1320-21 (quoting *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996)) (emphasis added).

Despite this rationale allowing for the use of implied definitions, the Federal Circuit recognized the problems inherent in relying on implied definitions. These problems include the potential of confining claims solely to the specific embodiments disclosed; importing limitations from

the specification into the claims; and difficulties in the interpretation of a term when the patentee acts as his own lexicographer. As a result, the *Phillips* Court acknowledged that there are no explicit rules governing whether the embodiments described in the specification merely provide examples or define the outer limits of the claim term.

Applying this approach to the construction of the claim term “baffles,” the en banc Court ruled that “baffles” can be mounted at any angle to the wall faces including right angles. 415 F.3d at 1327. To reach this conclusion, the majority stated that it based its decision on the claim language (using a claim differentiation theory) and on the specification. *Id.* Although the majority said that it found support in the specification for its interpretation, Judge Lourie disagreed. In his dissent, Judge Lourie noted that the specification refers to angled baffles “several times” and that the baffles are “disposed at such angles that bullets which might penetrate the outer steel panels are deflected.” *Id.* at 1329. Indeed, the dissent pointed out, the specification contains no disclosure of baffles at right angles. Thus, while stating that he would not “critique in detail” the majority’s statements made in “rationalizing” its claim construction holding, Judge Lourie concludes that it “could not be clearer” that baffles should be limited to those attached at an acute or obtuse angle. *Id.*

#### IV. POST-*PHILLIPS* JURISPRUDENCE AND THE REVIVAL OF IMPLICIT DEFINITIONS

Following *Phillips*, arguments have emerged regarding implicit definitions based on the specification based on the following factors: (1) when a patent makes consistent use of a term throughout the specification; (2) when the preferred embodiment is referred to as “the invention;” (3) when a structure is disclosed in the specification that, by its nature, exclude others; (4) when the specification distinguishes the prior art in a way that implies a claim limitation; (5) when the specification demans a technology or solution, and; (6) where the specification incorporates industry standards which in turn define a term. Each of these factors have guided decisions by various panels.

##### A. Consistent use of a term in the specification

If a patent makes repeated and consistent use of a term throughout the specification, certain panels of the Court have shown a substantially increased likelihood of looking to the specification to narrow the claim by interpretation. A good example of this analysis can be found in *Nystrom v. Trex Co., Inc.*, (“Nystrom II”) (Linn, Gajarsa, Mayer). 424 F.3d 1136 (Fed. Cir. 2005) *cert. denied*, 126 S. Ct. 1654 (2006). Nystrom, who held a patent for shaped flooring boards, brought an infringement action against a manufacturer of composite decking materials. The District Court construed the term “board” as used in the patent as limited to a board made from wood. In its first, though later superseded, decision written less than two months before *Phillips*, the Federal Circuit reversed in part and found the term referred to boards of any material. *Nystrom v. Trex Co., Inc.*, (“Nystrom II”), 374 F.3d 1105 (Fed. Cir. 2004). In reaching this decision, the Federal Circuit observed that nothing in the claims, dictionaries, or the specification limited the ordinary and customary meaning of board as encompassing any material. *Id.* at 1110-13.

The Federal Circuit panel granted Nystrom’s petition for rehearing, withdrew its initial opinion, and issued a replacement decision affirming the limitation of “board” to wood boards. Nystrom II, 374 F.3d 1145-46. The Court first considered the claim language and noted that the patent included dependent claims with the further limitation of a wood board. Although acknowledging that claim differentiation theory would create a presumption that the different meaning was intended, the Court concluded that “[d]ifferent terms or phrases in separate claims may be construed to cover the same subject matter where the written description and prosecution history indicate that such a reading of the terms or phrases is proper.” *Id.* at 1143 (quoting *Tandon Corp. v. United States Int’l Trade Comm’n*, 831 F.2d 1017, 1023-24 (Fed. Cir. 1987)).

The Court reached its conclusion in large part on the basis of the consistent use of “wood” as limiting board throughout the specification.<sup>2</sup> The Court noted that the specification repeatedly (but not exclusively) referred to a board as made of wood. The background of the invention, for example, “frames the invention in the context of wood decking materials cut from logs, even though it acknowledges that other materials exist,” and that this context is “maintained throughout the written description.” *Id.* at 1143. In that circumstance, even though the written description began by noting that a variety of materials have been used, the unanimous Court concluded that boards should be limited to those made of wood.

Similarly, in *Ocean Innovations, Inc. v. Archer* (Schall, Michel, Linn), the Court concluded that the claim term “floatation units” as used in a patent for a floating drive-on dry dock included a requirement that the units be hollow. 145 Fed. Appx. 366 (Fed. Cir. 2005). Citing *Phillips*, the Federal Circuit noted that “the line between construing terms and importing limitations can be discerned with reasonable certainty and predictability if the court’s focus remains on understanding how a person of ordinary skill in the art would understand the claim terms.” *Id.* at 370. Using this guideline, the Federal Circuit noted that first sentence of the patent’s abstract characterized the invention as being “assembled from two kinds of hollow floatation units,” that other references to hollow units were contained in the specification, and that the preferred embodiment states that all of the units are hollow. *Id.* These repeated references, together with a description of the invention with reference to prior art that contain hollow units, was sufficient to convince this panel to construe floatation units as hollow.

*Agfa Corp. v. Creo Products, Inc.* (Rader, Lourie, Newman dissenting) presents an interesting study of where the treatment of a term in the specification was not found to result in a limitation to a claim term. 451 F.3d 1366 (Fed. Cir. 2006). The majority in that case refused to limit the claim term “stack,” as it applied to a stack of printing plates, to a horizontal stack even though the preferred embodiment, and all descriptions of the invention, depicts stack as horizontal. *Id.* at 1376. In reaching its decision, the majority affirmed the District Court’s ruling and pre-*Phillips* reasoning.

The panel began by discussing with approval the trial court’s initial consultation of a dictionary for the meaning of “stack.” In doing so, the majority explained that stack has an “ordinary meaning” that “may be readily apparent even to lay judges, and claim construction in such cases involves little more than the application of widely accepted meaning of commonly understood words.” *Id.* at 1376 (citing *Phillips*, 415 F.3d at 1314). As a result, the majority concluded that the case fell “squarely within the guidelines from *Phillips*” of those situations in which general purpose dictionaries have been found to be helpful. *Id.* Next, the panel looked to the specification and noted that, while each and every depiction of a stack showed it in a horizontal arrangement, the specification otherwise provided no reference whatsoever to the orientation of the stack. Finally, the majority noted the existence of dependent claims that specify a horizontal arrangement to the stacks and, under a claim differentiation theory, concluded that “without any indication beyond the necessary depiction to suggest limiting the invention to this single embodiment, the broader language of the claims cannot carry the unexpressed and unintended (at the time of patent drafting) limitation.” *Id.* at 1377.

In her dissent, Judge Newman viewed the extent of the references to horizontal stacks in the specification differently. As she explained:

The claims and the entire descriptive text make clear that the invention is directed to ‘substantially horizontal’ plates that are stacked one upon the other. When the plates are stacked vertically – as in the prior art, and apparently in the

<sup>2</sup> The *Nystrom* Court also concluded that the prosecution history supported the wood limitation and reaffirmed the limited role of dictionaries. While the Court did not remove from its reissued opinion the panel’s review of the dictionary meaning of “board,” it reduced the relevance of such definitions well below the specification and located other dictionary definitions of board as made of wood. The Court then concluded that:

*Phillips* now counsels ...that in the absence of something in the written description and/or prosecution history to provide explicit or implicit notice to the public - i.e., those of ordinary skill in the art - that the inventor intended a disputed term to convey more than the ordinary and customary meaning revealed by the context of the intrinsic record, it is improper to read the term to encompass a broader definition simply because it may be found in a dictionary, treatise, or other extrinsic source. *Id.* at 1145.



accused system, there is no top plate. I cannot agree with the panel majority that “stack” is correctly construed to include the vertical alignment of the prior art.

*Id.* at 1384-85.

The juxtaposition of the panel and dissenting opinions in *Agfa* is interesting for several reasons. First, while Judge Newman saw the multiple drawings of a horizontal stack (and use of terms such as the “top” of the stack) as critical, the majority seems to have given the drawings less weight. Moreover, the majority’s initial reliance on the dictionary definition of “stack” is a departure from other post-*Phillips* cases that turn to dictionaries, if at all, only after a thorough review of the specification.

Even more recently, two of the judges who had signed-on to the *Nystrom II* opinion proved unwilling to use the specification to narrow the meaning of a claim term under slightly different circumstances. *Acumed LLC v. Stryker Corp.*, 483 F.3d 800 (Fed. Cir. 2007) (Gajarsa, Linn, Moore dissenting). In *Acumed*, the plaintiff’s patent covered a curved orthopedic nail used to pin together a fractured humerus. One of the contested terms was the “transverse holes” in the butt end of the nail, which receive screws to attach the device to a patient’s bone. The accused infringer argued unsuccessfully—both in the trial court and on appeal—that “transverse holes” should be limited to holes that are “perpendicular to the nail shaft.” *Id.* at 807.

Affirming the lower court’s interpretation of “transverse hole” as any hole “across the butt portion of the nail,” whether perpendicular to the shaft or not, Judge Gajarsa’s opinion began by acknowledging the infringer’s argument that “[e]very description of the transverse holes in the...patent contemplates a perpendicular hole.” *Id.* And, somewhat remarkably, the majority admitted that “[t]his is a correct characterization of the patent.” But citing the *Phillips* prohibition against confining patent claims to the preferred embodiment, the majority refused to limit “transverse holes” to perpendicular ones. *Id.* Then, for good measure, the majority opinion went on to parse the specification, insisting that the description of the preferred embodiment itself showed that there was a difference between transverse and perpendicular:

The written description states that Figure 2 ‘illustrates a plurality of transverse holes, each of which is...perpendicular to the portion of the nail axis at the butt portion...of the nail.’...This implies that a ‘transverse’ hole need not be ‘perpendicular’—if it were, the patentee would not have needed to clarify that these holes, in addition to being transverse, were perpendicular to the nail axis.

*Id.* at 807.

Relying in part on *Nystrom II*, Judge Moore’s dissent pointed out that “[t]here is not a single non-perpendicular, ‘transverse’ hole shown or described in the patent,” and argued that “[c]onstruing ‘transverse’ to include something other than perpendicular—in spite of the repeated, narrow usage of that term in the specification—would provide patent coverage that is broader than what the inventor actually invented and disclosed in his specification....” *Id.* at 814. In response to this dissent, the majority distinguished the result in *Nystrom II* by arguing that, in both cases, the court merely relied on the “ordinary” meaning of the disputed claim term. According to the majority, both parties in *Nystrom II* had agreed that the ordinary meaning of “board” was a piece of sawed lumber; all the court had done there was to refuse to *expand* the meaning of the term to cover other materials that were not contemplated by the specification. *Id.* at 808-809. And in *Acumed*, the majority indicated that it was simply applying the ordinary meaning of a “transverse hole” as a hole through some object, regardless of angle, and refusing to *narrow* the term based on the fact that all of the transverse holes described in the specification happened to be transverse.

## **B. Specification language identifying preferred embodiment as “the invention”**

Another circumstance in which some panels of the Federal Circuit have recently appeared more likely to imply a limitation from the specification is when the preferred embodiment is referred

to as “the invention.” Thus, post-*Phillips* cases looked back to *SciMed Life Systems, Inc. v. Advanced Cardiovascular Systems, Inc.*, for the proposition that “the characterization of [a limitation] as part of the ‘present invention’ is strong evidence that the claims should not be read to encompass the opposite structure.” 242 F.3d 1337, 1343 (Fed. Cir. 2001).

This result can be found in cases such as *IP Innovation L.L.C. v. Ecollege.com* (Rader, Newman, Dyk), 156 Fed. Appx. 317 (Fed. Cir. 2005), and *Honeywell International, Inc. v. ITT Industries, Inc.* (Lourie, Mayer, Dyk), 452 F.3d 1313 (Fed. Cir. 2006). In *IP Innovation*, for example, the claim at issue covered an interactive computer software system including pictures and graphics. The specification identified “the invention” as having certain attributes – for example control commands embedded in a single file in pictures. The panel concluded that, in that circumstance, it was appropriate to infer the definition of single file command into the claim.<sup>3</sup>

Similarly, the Court concluded in *Honeywell* that, where the specification referred to a fuel filter as “this invention” or “the present invention,” and where no components other than the fuel filter were described in the specification as composed of the required material, it was appropriate to limit the term “fuel injection system component” to a fuel filter. 452 F.3d at 1315.

The same reasoning was used by the court to limit claims covering “composite compositions” of wood and plastic in a group of four patents sharing substantially identical specifications in *Andersen Corp. v. Fiber Composites, LLC*, 474 F.3d 1361 (Fed. Cir. 2007) (Bryson, Prost, Saris). The Federal Circuit affirmed the lower court’s ruling that “composite compositions” would be limited to admixtures in the form of pellets or extrudates even though the at-issue claims did not carry that limitation on their face. *Id.* at 1365-66. Numerous statements in the specifications describing these forms as integral to ‘the invention’ led the court to conclude that “those statements are not descriptions of particular embodiments, but are characterizations directed to the invention as a whole,” meaning that they are indeed appropriately construed as limitations on the claims. *Id.* at 1367.

To be sure, there are cases reaching the opposite view. In *nCube Corp. v. Seachange International, Inc.* (Rader, Friedman, Dyk dissenting), for example, the plaintiff held a patent for a method of providing multimedia data in a networked system. 436 F.3d 1317 (Fed. Cir. 2006). The panel in that case declined to imply a limitation that logical (as opposed to physical) addresses be used to route messages even though the specification referred to the use of logical addresses as the scheme for “the present invention.” *Id.* at 1328. In reaching its conclusion, the majority made no reference to the use of the phrase “this invention” in the specification, focusing instead on the specification’s description of the invention as including divergence from the claimed embodiment. Judge Dyk dissented and took exception to the panel’s failure to give adequate weight to the specifications use of the phrase “the present invention”, explaining that he would limit the claims to the use of logical addresses.<sup>4</sup>

### C. Disclosure of structures that, by their nature, exclude alternatives

Some panels have implied a claim limitation based on the specification disclosing a structure which, by its very nature, excludes alternative structures and therefore bars a broader definition of the claim term.

The panel in *On Demand* (Newman, Mayer, Bryson), for example, limited “paper pages” to pages that have been pre-cut based on a structure disclosed in the specification. 442 F.3d 1331 (Fed. Cir. 2006). In that case, the patent was directed to a system and method for manufacturing a single book copy. The patentee sued a company that would create single copies of books at its factory, and Amazon, the company for whom the single book was created. The manufacturer printed the book on

<sup>3</sup> 156 Fed. Appx. at 322 (“Rather, the ‘404 patent’s narrow characterization of what ‘the invention’ covers in references to the embedded limitation supports a construction requiring location of all the command codes for a given picture or screen within a single file.”).

<sup>4</sup> *Id.* at 1328. See also *Paymaster Techs, Inc. v. United States*, 2006 WL 1228851, at \*7-8 (Fed. Cir. May 4, 2006) (Michel, Friedman, Dyk, J., dissenting).

rolls of paper that were subsequently cut and bound. The Federal Circuit reversed the District Court's interpretation of pages as encompassing rolls of paper thusly:

The specification states that the patented invention is the printing of individual copies of books at the customer site rather than in a "factory setting." . . . and illustrates a "page printer 26," which prints a "stack of paper text pages," showing page-size paper. The specification describes the Ross invention only as printing on individual pages, as distinguished from a "factory setting" and the production of multiple copies. *The Ross invention, and the claims, are directed to the on-site printing, and a binding of a single copy, for which printing on large webs and the requirement of cutting to page size would require equipment and procedures inimical to the substance of the Ross invention.*

*Id.* at 1342-43 (emphasis added).

Judge Dyk's dissent in *nCUBE Corp. v. SeaChange International, Inc.*, presents another example of finding a limitation based on an implicit definition arising from the very nature of the invention. 436 F.3d 1317 (Fed. Cir. 2006). As noted above, this case turned on whether a networked system routed various messages using only logical addresses, or could also use physical addresses. *Id.* at 1321. The majority of the panel concluded that either logical or physical addresses were permitted under a claim differentiation theory and because the specification "encompasses divergence from that embodiment." *Id.* Judge Dyk disagreed. After noting his view that the specification fails to provide a description of any invention beyond one using logical addresses, Judge Dyk stated "[i]ndeed, if the upstream manager could not route messages from the client using logical addresses, then the purpose of the invention—to substitute logical for physical addresses—would be defeated." *Id.* at 1328 (Dyk, J., dissenting).

Another case finding a limitation based on a specification's exclusion of alternative structures is *DSU Med. Corp. v. JMS Co., LTD.*, 471 F.3d 1293 (Fed. Cir. 2006). In a case dealing with patents for winged needle assemblies designed to reduce the risk of accidental needle pricks, the court in DSU construed the term "slidably enclosing" to require "that the guard substantially contain the needle-assembly at all times." *Id.* at 1300. The trial court then determined that because the allegedly infringing product was a stand-alone guard without a needle, there was no infringement. *Id.* The court of appeals found that the language and context of the claims supported the trial court's construction. The claim used the term "enclosing" in the context of an invention "for locking a needle in a shielded position as the needle is removed from a patient." *Id.* Additionally, the court of appeals found that the specification and the figures contained in the specification reinforced this construction and conveyed the concept of a permanent cover for the needle. *Id.*

In *Akeva L.L.C. v. Addidas-Saloman*, 208 Fed. Appx. 861 (Fed. Cir. 2006), Akeva sued Addidas for infringement of certain claims of two of its patents for improving athletic shoe heels with "detachable, or rotatable, or permanently secured" rear heels. *Id.* at 862. The district court construed the patents as covering shoes with rotatable or detachable rear soles only, and because Addidas shoes had fixed, non-rotating heels, the court granted Addidas' motion of summary judgment for non-infringement. The court of appeals affirmed. The court noted that "[w]here the specification makes clear that the invention does not include a particular feature, that feature is deemed to be outside the reach of the claims of the patent, even though the language of the claims, read without reference to the specification, might be considered broad enough to encompass the feature in question." *Id.* at 865 (quoting *SciMed Life Sys. Inc. v. Advanced Cardiovascular Sys. Inc.*, 242 F.3d 1337, 1341 (Fed. Cir. 2001)).

For the '300 patent at issue, the district court construed the term "secured" to mean "selectively or permanently fastened, but not permanently-fixed into position." *Id.* at 864. The court found that the "permanently attached" language in the specification contemplated shoes with heels that are permanently fixed, but rotatable. *Id.* at 865. Additionally, the specification mentioned that a problem with prior art was that a shoe owner would be "stuck" with their original heels. *Id.* The court

found that nothing in the specification contemplated a rear sole that would be permanently locked into position. *Id.* As such, the term “secured” had to be limited to shoes with rear soles that are secured to the shoe, but not permanently fixed. *Id.*

#### D. Denigration of technologies or solutions in the specification

Implicit definitions have also been applied where the court finds an implicit exclusion, for example where the specification demeans an alternative solution, or where the specification provides limiting implicit context.

This approach can be seen most clearly in *Honeywell International, Inc. v. ITT Industries, Inc.*, (Lourie, Mayer, Dyk), 452 F.3d 1312 (Fed. Cir. 2006). In addition to the claim construction issues discussed above, that case involved the construction of “electrically conductive fibers.” Although the patent holder argued, and the District Court agreed, that the term should be broadly interpreted to include fibers made of both metal and carbon, the Federal Circuit reversed, and excluded carbon fibers from the claim. The Federal Circuit reached this conclusion because the specification touted the advantages of stainless steel fibers and, indeed, demeaned carbon filters.<sup>5</sup> Finding that “the written description has gone beyond expressing the patentee’s preference for one material over another” the Court concluded that patent holder’s “repeated derogatory statements concerning one type of material are the equivalent of disavowal of the subject matter from the scope of the patent’s claims.” *Id.* at 1320.

Similarly, the Court in *Wireless Agents LLC v. Sony Ericsson Mobile Communications AB* (Dyk, Gajarsa, Rader) found persuasive the denigration of an alternative solution in the specification when imposing a limitation from the specification. That case involved interpretation of the term “alphanumeric keyboard” in the context of a patent covering a handheld electronic communication device. 2006 WL 2091197 (Fed. Cir. July 26, 2006). The Federal Circuit construed an “alphanumeric keyboard” as requiring a substantially full set of alphabetic and numeric keys, and excluded from its construction a twelve-digit telephone keypad. *Id.* at \*2. In reaching that conclusion, the Court noted that, while the meaning of the disputed term was not apparent from the face of the claim, the specification explicitly described the disadvantages of keypads that have only twelve keys as “extremely slow, awkward, error prone, and not appropriate for a device intended to transfer textual data on a regular basis . . . .” *Id.* The specification further distinguished the “alphanumeric keyboard” as “easier and faster to use and learn” than common mobile phone keypads. *Id.*

This was not the case in *Ventana Med. Sys. Inc. v. BioGenex Lab. Inc.*, 473 F.3d 1173 (Fed. Cir. 2007), *rehearing en banc denied* (Feb. 7, 2007). In *Ventana*, the district court construed the term “dispensing” to require direct dispensing, meaning that “the regent is dispensed directly from the regent container onto the slide rather than utilizing an intermediate transport mechanism to transfer the regent from the regent container to the slide.” *Id.* at 1178. *Ventana* appealed the district court’s finding of non-infringement and the sole issue on appeal was the proper construction of the term “dispensing.” The court of appeals declined to limit the term “dispensing” to “direct dispensing.” The court found that *Ventana*’s general statements about previously known devices in the BACKGROUND ART section of its patent did not constitute a disavowal of coverage by *Ventana* over features in the prior art. *Id.* at 1180-1181. Unlike the case in *Honeywell*, *BioGenex* was only able to point to general statements indicating that the invention was intended to improve upon prior art automated staining methods and such general statements, without more, cannot be “interpreted to disclaim every feature of every prior art device discussed in the BACKGROUND ART section of the patent.” *Id.* at 1181.

Also of interest in this regard is *Wilson Sporting Goods Co. v. Hillerich & Bradsby Co.* (Rader, Lourie, Bryson), which reversed the District Court’s interpretation of a softball bat’s “insert” by adding the implied definition that the insert be “rigid.” 442 F.3d 1322 (Fed. Cir. 2006). The trial court construed “insert” to mean “a rigid, circular, hollow tube having an outer diameter less than the

---

in part).

5 *Id.* at 1319-20 (“Unlike the non-metallic fibers, stainless steel fibers also do not increase mechanical strength or stiffness of the base resin significantly.

inner diameter of the tubular frame impact portion.” *Id.* at 1325. The Federal Circuit disagreed, stating that “[n]othing in the claims or specification indicates, explicitly or implicitly, that the inventor intended to impart a novel meaning to ‘insert.’” *Id.* at 1330. What seemed to be missing from the specification in *Wilson* (in the eyes of the Federal Circuit but not the District Court), that was present in the specifications in other cases, is a suggestion that a non-rigid insert would be inferior (or that a rigid insert would be an advance in the technology). 442 F.3d at 1330.

### **E. Method of distinguishing prior art that favor the imposition of implied definitions**

Implied definitions have also been found where the patentee distinguishes the patent over prior art in a way that evidences a limiting claim definition. This approach can be seen in *Honeywell v. ITT*, 452 F.3d 1312. As noted above, the Federal Circuit affirmed the District Court’s limitation of “fuel system component” to a fuel filter in connection with a patent covering an electrostatically dissipative fuel system. In support of its ruling, the Federal Circuit reviewed the specification’s discussion of the technological problem that the invention was trying to solve in an advance over the prior art. The prior art covered fuel filters made of a polymer material. With the advent of electronic fuel injection systems, microscopic holes were created in a polymer fuel filter when electrostatic charges built up and then discharged by arcing through the fuel filter to the metal frame of a car. The patent covered the incorporation of electrostatically conductive material into the polymer housing of the device. *Id.* at 1321.

The *On Demand v. Ingram* case discussed above provides another example. 442 F.3d 1331. In that case, the Court focused on the manner in which the invention distinguishes the prior art when it ordered a limitation (that a “customer” must be an individual consumer) based on an implied definition found in the specification. In reaching its conclusion, the Court reasoned that the prior art covered the method of electronically taking orders for books, and the printing of a single copy of a book. Noting that in both the specification and during reexamination the patent holder distinguished the prior art from the invention, the Court concluded that the combination of the references, when considered in connection with the language of the specification, established that the “direct customer role contemplated in the Ross patent...is central to the claim construction.” *Id.* at 1345.

The ruling in *Old Town Canoe Co. v. Congluence Holdings Corp.* (Linn, Schall, Mayer) is interesting in that it implies a limitation based not on the way that the invention is distinguished from the prior art, but on how it is similar. 448 F.3d 1309 (Fed. Cir. 2006). *Old Town Canoe* involved a patent covering a method to make plastic laminate boat hulls. The panel agreed with the defendant that the claim term “coalescence,” as used in connection with a multistep process for creating a boat hull, includes a limitation that the process be continued until it reaches its optimal state. In reaching this conclusion, the Court noted that the written description described the invention as an improvement over a patented coalescence process that operated in a different way but also reaches an optimal state. *Id.* at 1317. The *Old Town Canoe* panel concluded that the prior art cited in the specification supported the conclusion that one of ordinary skill in the art would understand “coalescence” to refer to a process that proceeded to an optimal state. *Id.*

On the other hand, *Ventana* is a good example of the Court’s declining to narrow the claim definition based on the fact that the specification distinguished prior art. As discussed above, *Ventana* dealt with whether the term “dispensing” was limited to “direct dispensing.” 473 F.3d at 1180. The Appellee, Biogenex, contended that the specification had expressly distinguished prior art including the “sip and split” method and therefore the invention at issue could not practice that method of dispensing. *Id.* The Court, however, found that the specification also listed prior art that employed a “direct dispensing” method, and therefore, Biogenex’s argument would lead to a nonsensical result where the patent’s preferred method would also be excluded from the definition. *Id.* As such, the Court declined to narrow the term. *Id.* Industry standards as implied definitions.

An intriguing scenario in which claims have been interpreted in light of a specification and a narrowing implied definition found is that described in *LG Electronics, Inc. v. Bizcom Electronics, Inc.* (Mayer, Michel, Newman). 453 F.3d 1364 (Fed. Cir. 2006). That case involved a patent disclosing a digital computer system which interconnects devices using a system bus. In interpreting the term

“requesting agent,” the patent holder argued that an industry standard, which was incorporated by reference into the specification, supplied the correct definition of “requesting agent” as “an agent that has entered into the arbitration function for bus access.” *Id.* at 1375. The Federal Circuit agreed, overturning the decision of the District Court.

In so ruling, however, the Federal Circuit explained that incorporation of an industry standard does not automatically supply the correct definition. *Id.* at 1374. Rather, the Court reasoned that an industry standard is highly relevant—and is indeed intrinsic evidence—in determining the meaning of the claim term to one of ordinary skill in the art. Giving the industry standard its due weight in determining claim construction, the Court ruled that the industry standard should imply the proper definition of the disputed term in the claim. *Id.* at 1375.

## F. The Persistent Factor of the “Sole Embodiment”

The *Phillips* decision was clear: the fact that a specification discloses a single embodiment does not provide the basis for reading a limitation into the claim. 415 F.3d at 1323. Subsequent panels have reaffirmed this principle. In *JVW Enterprises, Inc. v. Interact Accessories, Inc.*, for example, the panel found improper the defendant’s request to limit claim language to the sole embodiment. 424 F.3d 1324 (Fed. Cir. 2005) (Prost, Michel, Schall). In reaching its decision, the panel reiterated the *Phillips* rule that “even when a specification describes very specific embodiments of the invention or even describes a single embodiment” the court should not “import limitations into the claims from examples or embodiments appearing only in a patent’s written description.” *Id.* at 1335. The only exception to this, the *JVW* panel noticed is if “the specification makes clear that the patentee . . . intends for the claims and the embodiments in the specification to be strictly coextensive.” *Id.*

Nonetheless, the impact of the fact that a specification has a single embodiment persists. Thus, some panels seem eager to limit the reach of the claim terms by focusing on the written description requirement. In *LizardTech, Inc. v. Earth Resources Mapping, Inc.*, (Bryson, Lourie, Schall), for example, the Court limited a claim to the only disclosed embodiment by finding that anything beyond that embodiment would fail to meet the requirements of section 112. 424 F.3d 1336 (Fed. Cir. 2006). The *LizardTech* panel reviewed the requirements that a patent must describe the invention such that one skilled in the art can make and use it and that it must describe the invention sufficiently to give notice that the patentee had invented what is claimed. Considering these together, the Court concluded that construing a claim as broader than its description would result in a patent with a claim scope “far greater than what a person of skill in the art would understand the inventor to possess or what a person of skill in the art would be enabled to make and use.” *Id.* at 1346.

In his dissent to the denial of a petition for rehearing of the *LizardTech* case en banc, Judge Rader also explained his disinclination to limit claims to the embodiment contained in the specification. Judge Rader’s dissent provides a detailed review of the sole embodiment doctrine and noted the tension between *JVW*, in which a panel of the Federal Circuit concluded that the claims should not be narrowed to the preferred embodiment “unless the specification suggest the inventor intended such narrow coverage,” and the panel’s ruling in *LizardTech* that claims should not be broader than the disclosed embodiments “unless the specification suggests the invention is broader than those embodiments.” *LizardTech*, 433 F.3d 1373, 1376-77 (Fed. Cir. 2005). Ultimately reiterating the binding nature of the en banc ruling on sole embodiment in *Phillips*, Judge Rader opined that “the written description analysis of *LizardTech* is troubling, if not inexplicable.” *Id.* at 1377-78.

*Inpro II Licensing, S.A.R.L. v. T-Mobile USA, Inc.* (Newman, Dyk, Prost) suggested more direct resistance to the established rule. The patent in *Inpro* covered modules for a personal digital assistant (PDA) that controlled the transfer of data between a PDA and a host computer. The Court construed “host interface” as a “direct bus interface” in large part because that was the only interface described in the specification. 450 F.3d 1350, 1355 (Fed. Cir. 2006). Recognizing that “claims need not be limited to the preferred embodiment, when the invention is more broadly described,” the Court concluded that “neither do the claims enlarge what is patented beyond what the inventor has described as the invention.” *Id.* (quoting *Netword, LLC v. Centraal Corp.*, 242 F.3d 1347, 1352 (Fed. Cir. 2001)).

This point was similarly addressed by the Court in *On Demand v. Ingram*, 442 F.3d 1331. As noted above, the *On Demand* panel inferred into the definition of “customer” the requirement found in the specification that the customer be one that orders and receives a single-copy printing at the customer’s site. In reaching this conclusion, the panel explained “[i]n *Phillips*, 415 F.3d at 1321, the *en banc* court explained that the role of the specification is to describe and enable the invention. In turn, the claims cannot be of a broader scope than the invention that is set forth in the specification.” *Id.* at 1340.

## V. CONCLUSION

A review of the Federal Circuit’s post-*Phillips* claim construction reveals that certain principles may be emerging to help determine when an implicit definition will be found based on the specification. Many of these approaches were in use pre-*Phillips*. Others seem to have gained momentum since the *en banc* decision. In any event, a close reading of the Federal Circuit’s recent jurisprudence reveals themes that a litigant could invoke (if the facts of the individual case warrant) to support an implied definition of a claim term.

# SOLVING HOBSON'S CHOICE: SUGGESTED CHANGES TO WILLFULNESS LAW IN THE WAKE OF *KNORR-BREMSE* AND *ECHOSTAR*

---

*Timothy J. Malloy & Merle S. Elliott  
McAndrews, Held & Malloy  
Chicago, IL*

Defendants accused of willful patent infringement are currently put to a Hobson's choice: waive attorney-client privilege, or risk a jury verdict of willfulness and potentially trebled damages.

In nearly every patent case, defendants are accused of willful infringement.<sup>1</sup> One of the frequently employed ways for a defendant to combat a willfulness charge is to rely on the advice of counsel. This "solution," however, results in a waiver of the attorney-client privilege regarding that advice.

The Federal Circuit recently weighed in on the issue of willfulness and the defense of advice of counsel in *Knorr-Bremse Systeme Fuer Nutzfahrzeuge GmbH v. Dana Corp.*<sup>2</sup> and *In re EchoStar Communications Corp.*<sup>3</sup> In *Knorr-Bremse*, the Federal Circuit reaffirmed that when a potential infringer has "actual notice" of a patent, the potential infringer has "an affirmative duty to exercise due care to determine whether or not he is infringing, including the duty to seek and obtain competent legal advice from counsel before the initiation of any possible infringing activity."<sup>4</sup> The Federal Circuit also partially alleviated the traditional defendant's dilemma, holding that it is the patent holder's burden to prove willful infringement, and where a defendant asserts the attorney-client privilege and withholds the advice of counsel, there will be no "adverse inference as to the nature of the advice."<sup>5</sup>

Unfortunately, the Federal Circuit's opinion in *Knorr-Bremse*, especially in light of the more recent *EchoStar* opinion, does not end the defendant's problem. Many defendants will still choose to rely on the advice of counsel, and the scope of the waiver of privilege, as discussed in *EchoStar*, has been interpreted so broadly by some courts that it may jeopardize a defendant's ability to receive effective representation.

The defendant's dilemma regarding willfulness charges and the advice of counsel has not escaped the notice of Congress. In the recently-introduced Patent Reform Act of 2007, Congress is considering amendments to the patent law that would substantively change the treatment of willfulness charges. In this proposed legislation, a charge of willfulness cannot be pled until after a determination that a patent is infringed and not invalid. Willfulness would also be made an issue for the court.

1 Kimberly A. Moore, *Empirical Statistics on Willful Patent Infringement*, 14 FED. CIR. B.J. 227, 232 (2004) (patentees charged defendants with willful infringement in more than 90 percent of all cases; data discussed in a section titled "Willfulness is Alleged All the Time").

2 *Knorr-Bremse Systeme Fuer Nutzfahrzeuge GmbH v. Dana Corp.*, 383 F.3d 1337 (Fed. Cir. 2004) (en banc).

3 *In re EchoStar Communications Corp.*, 448 F.3d 1294 (Fed. Cir. 2006).

4 *Underwater Devices, Inc. v. Morrison-Knudson Co.*, 717 F.2d 1380, 1389-90 (Fed. Cir. 1983).

5 *Knorr-Bremse*, 383 F.3d at 1345.



Should Congress take no action regarding willfulness, however, there are at least two potential solutions to the defendant's dilemma available after the recent Federal Circuit decisions in *Knorr-Bremse* and *EchoStar*. First, courts could exercise the case management powers they already have to bifurcate trial – and discovery – on willfulness and advice of counsel issues until after a finding on liability. Second, the Federal Circuit could reconsider its position that willfulness is a question of fact, to be determined by the jury. This paper suggests consideration as to whether willfulness is more appropriately an equitable issue, analogous to inequitable conduct, and should be determined by the court.

## I. TWO UNSAVORY OPTIONS: THE CURRENT STATE OF WILLFULNESS LITIGATION

Whether infringement is “willful” is currently a question of fact to be determined by the trier of fact (normally a jury).<sup>6</sup> A determination of willfulness is made by considering the totality of the circumstances, including the factors enumerated in *Rolls-Royce Ltd. v. GTE Valeron Corp.* and *Read Corp. v. Portec, Inc.*<sup>7</sup> If infringement is found to be willful, the court can then treble the defendant's damages liability and award the plaintiff its attorneys' fees.<sup>8</sup>

Defendants faced with a charge of willful infringement thus must make a distasteful decision: waive the attorney-client privilege or risk a potentially undeserved finding of willfulness and trebled damages. The problem is magnified by the amounts often at stake.

On one hand, defendants are often reluctant to waive the attorney-client privilege for good reason - waiving the privilege erodes the freedom of attorney-client communication, and therefore the effectiveness of counsel. Yet, a finding of willfulness and trebled damages entered against a defendant are also undesirable to say the least.

Adding further complexity to the defendant's dilemma, the defendant is often forced to choose whether to rely on the advice of counsel as a defense to a willfulness charge before any finding is entered on liability, and sometimes before significant discovery on liability issues has been conducted.

Recent Federal Circuit decisions in the area of willfulness, including *Knorr-Bremse* and *EchoStar*, do not eliminate this dilemma. While the Federal Circuit in *Knorr-Bremse* removed any adverse inference drawn from the failure of the defendant to obtain or produce an opinion of counsel, it still maintained the defendant's “duty of care” to avoid infringing on others' patents. In *EchoStar*, the Federal Circuit addressed the scope of the waiver of the attorney-client privilege and attorney work product where a defendant does choose to rely on the advice of counsel.

Based on the decisions in *Knorr-Bremse* and *EchoStar*, as well as the way those decisions have been interpreted and applied by district courts, forcing a defendant to choose whether or not to waive privilege at an early stage of the litigation erodes, if not eliminates, a defendant's ability to obtain effective representation.

### A. A Small Step In The Right Direction: *Knorr-Bremse*

As discussed above, willfulness is a question of fact, based on consideration of the totality of the circumstances. In *Read Corp. v. Portec, Inc.*, the Federal Circuit enumerated factors adopted in *Knorr-Bremse* to be considered in a willfulness determination.<sup>9</sup> These factors include: (1) whether the infringer deliberately copied the ideas or design of another; (2) whether the infringer, when he knew of the other's patent protection, investigated the scope of the patent and formed a good-faith belief that it was invalid or that it was not infringed; (3) the infringer's behavior as a party to the litigation; (4) defendant's size and financial condition; (5) closeness of the case; (6) duration of defendant's

<sup>6</sup> *Knorr-Bremse*, 383 F.3d at 1342-43.

<sup>7</sup> *Id.* (citing *Rolls-Royce Ltd. v. GTE Valeron Corp.*, 800 F.2d 1101, 1110 (Fed. Cir. 1986), and *Read Corp. v. Portec, Inc.*, 970 F.2d 816, 826-27 (Fed. Cir. 1992)).

<sup>8</sup> 35 U.S.C. Section 284 (“the court may increase the damages up to three times the amount found or assessed”); 35 U.S.C. Section 285 (“the court in exceptional cases may award reasonable attorney fees to the prevailing party”).

<sup>9</sup> *Knorr-Bremse*, 383 F.3d at 1342-43.

misconduct; (7) remedial action by the defendant; (8) defendant's motivation for harm; and (9) whether defendant attempted to conceal its misconduct.<sup>10</sup> A defendant often obtains and relies on the advice of counsel in order to develop a "good-faith belief that [a patent] was invalid or that it was not infringed."<sup>11</sup>

Where, however, a defendant had actual notice of a patent but did not provide or rely on the advice of counsel, the Federal Circuit drew an "adverse inference." In *Kloster Speedsteel AB v. Crucible Inc.*, the Federal Circuit some twenty years ago held that the infringer's failure to produce an opinion, "in alleged reliance on the attorney-client privilege, would warrant the conclusion that it either obtained no advice of counsel or did so and was advised that its importation and sale of the accused products would be an infringement of valid U.S. patents."<sup>12</sup> The Federal Circuit affirmed the adverse inference in later cases, establishing a general rule that "a court must be free to infer that either no opinion was obtained or, if an opinion were obtained, it was contrary to the infringer's desire to initiate or continue its use of the patentee's invention."<sup>13</sup>

Recognizing that application of the adverse inference "resulted in inappropriate burdens on the attorney-client relationship," the Federal Circuit took the *Knorr-Bremse* case *en banc* in order to reconsider its precedent regarding the adverse inference.<sup>14</sup> The Federal Circuit, in overruling its precedent authorizing an adverse inference, explained that the "conceptual underpinnings" of this precedent have significantly diminished in force.<sup>15</sup> The advice of counsel defense and attendant waiver of the attorney client privilege spawned what the Federal Circuit described as "extensive satellite litigation."<sup>16</sup>

The Federal Circuit in *Knorr-Bremse* addressed four questions regarding the adverse inference. Specifically, the Federal Circuit held: (1) the assertion of the attorney/client privilege and the withholding of the advice of counsel shall no longer entail an adverse inference as to the nature of the advice; (2) the failure to obtain an opinion of counsel will not entail an adverse inference or evidentiary presumption that such an opinion would have been unfavorable; (3) in the context of the facts in this particular case, a fresh weighing of the evidence was required to determine whether the defendant committed willful infringement; and (4) the existence of a substantial defense to infringement is relevant but not *per se* sufficient to defeat liability for willful infringement.<sup>17</sup>

In abolishing the adverse inference where an opinion is obtained but the attorney-client privilege is asserted, the Federal Circuit emphasized that the public benefits when clients, in this case defendants, feel confident in sharing information with attorneys. "There should be no risk of liability in disclosures to and from counsel in patent matters; such risk can intrude upon full communication and ultimately the public interest in encouraging open and confident relationships between client and attorney."<sup>18</sup>

In abolishing the adverse inference where no opinion was obtained, the Federal Circuit did not consider the attorney-client privilege, but whether there is a "legal duty" upon a defendant to consult with counsel.<sup>19</sup> The Federal Circuit recognized the burden and cost of requiring "early and full study" of virtually every potentially adverse patent of which a company has knowledge.<sup>20</sup> The Federal Circuit maintained, however, that "there continues to be 'an affirmative duty of due care to avoid infringement of the known patent rights of others.'"<sup>21</sup>

Abolishing the adverse inference is a step in the right direction – it limits the penalty for a defendant who chooses to assert the attorney-client privilege or who chooses not to get an opinion.

10 *Read Corp.*, 970 F.2d at 826-27.

11 *Id.*

12 *Kloster Speedsteel AB v. Crucible Inc.*, 793 F.2d 1565, 1580 (Fed. Cir. 1986).

13 *Fronson v. W. Litho Plate & Supply Co.*, 853 F.2d 1568, 1572-73 (Fed. Cir. 1988).

14 *Knorr-Bremse*, 383 F.3d at 1343-44.

15 *Id.* at 1344 (citation omitted).

16 *Id.* at 1345.

17 *Id.* at 1345-47.

18 *Id.* at 1344.

19 *Id.* at 1345.

20 *Id.*

21 *Id.* at 1345-46 (quoting *L.A. Gear Inc. v. Thom McAn Shoe Co.*, 988 F.2d 1117, 1127 (Fed. Cir. 1993)).

Abolishing the adverse inference does not completely solve the defendant's dilemma, however. The Federal Circuit maintained the affirmative duty of due care,<sup>22</sup> so patent holders may argue that a defendant who has knowledge of a patent should have a difficult time showing that it exercised due care without presenting an opinion of counsel. As Judge Dyk recognized in his partial dissent in *Knorr-Bremse*, "the majority opinion does not address whether a potential infringer can satisfy the requirement of due care without securing and disclosing an opinion of counsel."<sup>23</sup>

In addition, the Federal Circuit specifically failed to address the question of whether the existence or lack thereof of an opinion is appropriate evidence to present to the trier of fact. The Federal Circuit specifically stated that it was not ruling that the existence of opinions of counsel was irrelevant:

Several amici curiae raised the question of whether the trier of fact, particularly a jury, can or should be told whether or not counsel was consulted (albeit without any inference as to the nature of the advice received) as part of the totality of the circumstances relevant to the question of willful infringement. The amici pointed to various hypothetical circumstances in which such information could be relevant, even when there was no issue of attorney-client privilege. That aspect is not raised by this case, was not before the district court, and has not been briefed on this appeal. Today we resolve only the question of whether adverse inferences of unfavorable opinions can be drawn, and hold that they can not.<sup>24</sup>

*Knorr-Bremse* abolished the "adverse inference," but did not explicitly abolish the relevance of opinions of counsel apart from any inference.<sup>25</sup> A jury, if told that a defendant knew about a patent and obtained an opinion of counsel that it refuses to disclose, might well draw an adverse inference regarding that opinion on its own, despite any contrary instruction from a court.

The Federal Circuit's decision in *Knorr-Bremse* acknowledges that a defendant facing a willfulness charge must make a difficult decision regarding the advice of counsel and the attorney-client privilege, and takes at least a step towards providing a framework that is fair for a defendant. Even after *Knorr-Bremse*, however, a defendant may decide, from a practical standpoint, to produce an opinion of counsel in order to defend against a willfulness charge. Where a defendant does choose to waive privilege, the Federal Circuit's opinion in *EchoStar* is of little comfort regarding the scope of that waiver.

## B. A Step Back - *EchoStar*

In *In re EchoStar Communications Corp.*, the Federal Circuit recently addressed the scope of the waiver of the attorney-client privilege and the attorney work product immunity where the advice of counsel defense is asserted.<sup>26</sup> After *EchoStar* and its progeny, the scope of the waiver is uncertain, and may be so broad as to essentially deprive a defendant of effective counsel.

In *EchoStar*, the plaintiff sought discovery regarding the opinion on infringement provided to defendant by outside counsel after the filing of the lawsuit.<sup>27</sup> The Federal Circuit stated that a defendant should not be able to use the attorney-client privilege and the work product immunity "as both a sword and a shield," and recognized that "when a party defends its actions by disclosing an attorney-client communication, it waives the attorney-client privilege as to all such communications regarding the same subject matter."<sup>28</sup> Therefore, one could argue that if a defendant has received two opinions regarding the patent in suit, one favorable and one less favorable, in order to rely on the favorable opinion, the defendant must waive the privilege as to both.

22 In his partial dissent in *Knorr-Bremse*, Judge Dyk argued that enhancing damages for mere failure to exercise "due care" is inconsistent with Supreme Court jurisprudence holding that "punitive damages can only be awarded in situations where the conduct is reprehensible." *Id.* at 1348 (Dyk, J., dissenting-in-part).

23 *Id.* at 1352 (Dyk, J., dissenting-in-part).

24 *Id.* at 1346-47.

25 See, e.g., Joshua Stowell, *Willful Infringement and the Evidentiary Value of Opinion Letters After Knorr-Bremse v. Dana*, 2005 DUKE L. & TECH. REV. 5 (2005); Arthur Beeman, *Patent Opinions*, DAILY JOURNAL (2005), available at [http://www.lawinternational.com/articles/pub\\_146.pdf](http://www.lawinternational.com/articles/pub_146.pdf).

26 *In re EchoStar Communications Corp.*, 448 F.3d 1294 (Fed. Cir. 2006).

27 *Id.* at 1297.

28 *Id.* at 1301.

The Federal Circuit's decision in *EchoStar* also clarifies the scope of the waiver, largely as it applies to work product, by considering documents in three categories. First, those "documents that embody a communication between the attorney and client concerning the subject matter of the case, such as a traditional opinion letter," are included within the scope of the waiver.<sup>29</sup>

Second, documents that "discuss a communication between attorney and client concerning the subject matter of the case but are not themselves communications to or from the client," are included within the scope of the waiver.<sup>30</sup> Documents that reference a communication between attorney and client, even though not communicated to the client, do fall within the scope of the waiver. While the Federal Circuit recognized the danger that these documents may also contain legal analysis that was not communicated (and should be redacted), these documents "will aid the parties in determining what communications were made to the client and protect against intentional or unintentional withholding of attorney-client communications from the court."<sup>31</sup>

Third, documents that "analyz[e] the law, facts, trial strategy, and so forth that reflect the attorney's mental impressions but were not given to the client" do not fall within the scope of the waiver.<sup>32</sup> Because the "work-product waiver extends only so far as to inform the court of the infringer's state of mind," documents that contain attorney analysis that was not communicated to the client do not fall within the scope of the waiver.<sup>33</sup> "Counsel's opinion is not important for its legal correctness . . . . It is what the alleged infringer knew or believed, and by contradistinction not what other items counsel may have prepared but did not communicate to the client, that informs the court of an infringer's willfulness."<sup>34</sup>

While *EchoStar* clarifies to some extent the specific documents that should be produced when a defendant chooses to waive the attorney-client privilege and rely on the advice of counsel as a defense to a willfulness charge, *EchoStar* leaves larger questions unanswered: Does the waiver apply to trial counsel as well as opinion counsel? Does the waiver apply to all defenses, including infringement, invalidity, and unenforceability rather than just the scope of the particular opinion at issue?

For example, *EchoStar* discusses whether the waiver applies to advice given after litigation has commenced, including, potentially, advice from trial counsel rather than simply opinion counsel:

*EchoStar* contends that waiver of opinions does not extend to advice and work product given after litigation began. While this may be true when the work product is never communicated to the client, it is not the case when the advice is relevant to ongoing willful infringement, so long as that ongoing infringement is at issue in the litigation.<sup>35</sup>

Regarding the scope of subject matter waived, the *EchoStar* opinion is similarly unclear. The court states that "when a party defends its action by disclosing an attorney-client communication, it waives the attorney-client privilege as to all such communications regarding the same subject matter."<sup>36</sup>

In the brief wake of *EchoStar*, district courts have recognized these unanswered questions and reached differing conclusions.<sup>37</sup>

<sup>29</sup> *Id.* at 1302.

<sup>30</sup> *Id.*

<sup>31</sup> *Id.* at 1304.

<sup>32</sup> *Id.* at 1302.

<sup>33</sup> *Id.* at 1303 (emphasis omitted).

<sup>34</sup> *Id.*

<sup>35</sup> *Id.* at 1302 n.4.

<sup>36</sup> *Id.* at 1301.

<sup>37</sup> Prior to *EchoStar*, numerous district courts addressed the issue of the scope of the waiver of attorney-client privilege and work product immunity when a defendant relies on the advice of counsel as a defense to a willfulness charge. *See, e.g., Akeva LLC v. Mizuno Corp.*, 243 F. Supp. 2d 418 (M.D.N.C. 2003); *Motorola, Inc. v. Vasi Tech., Inc.*, No. 01-C-4182, 2002 U.S. Dist. LEXIS 15655 (N.D. Ill. Aug. 19, 2002); *Thorn EMI N. Am., Inc. v. Micron Tech., Inc.*, 837 F. Supp. 616 (D. Del. 1993); *Mushroom Assocs. v. Monterey Mushrooms, Inc.*, 24 U.S.P.Q.2d 1767 (N.D. Cal. 1992). These decisions are not discussed here partially because they pre-date *EchoStar*, and also because they have been amply discussed elsewhere. *See, e.g.,* William F. Lee & Lawrence P. Cogswell, III, *Understanding and Addressing the Unfair Dilemma Created by the Doctrine of Willful Patent Infringement*, 41 Hous. L. Rev. 393 (2004); Mark A. Lemley and Ragesh K. Tangri, *Ending Patent Law's Willfulness Game*, 18 BERKELEY TECH. L.J. 1085, 1100 (2003); David O. Taylor, *Wasting Resources: Reinventing the Scope of Waiver Resulting from the Advice-of-Counsel Defense to a Charge of Willful Patent Infringement*, 12 TELL. PROP. L.J. 319 (2004). Further, at least one district court has concluded that "the Federal Circuit's opinion in *Echostar* is not binding on this court." *Addas Am., Inc. v. Payless Shoesource, Inc.*, No. CV 01-1655-RE, 2006 U.S. Dist. LEXIS 79154, at \*8 (D. Or. Oct. 19, 2006).

The Federal Circuit may also weigh in again soon: In *In re Seagate Technology, LLC*, Misc. Docket No. 830 (Fed. Cir. 2007), the Federal Circuit invited the parties to brief the question of whether the waiver of attorney-client privilege and attorney work product extends to trial counsel as well as opinion counsel.<sup>38</sup> It is impossible to predict whether any Federal Circuit decision in *In re Seagate* would answer any or all of the remaining questions regarding waiver.

### 1. Does Waiver Extend to Trial Counsel?

Several district courts have held that the waiver of privilege and work product immunity does not extend to trial counsel. In *Indiana Mills & Manufacturing, Inc. v. Dorel Industries, Inc.*, the district court held that the temporal scope of the waiver did not extend past the filing of the lawsuit.<sup>39</sup> Although recognizing the holding of the Federal Circuit in *EchoStar* that an opinion provided after the filing of the lawsuit was discoverable, the district court in *Indiana Mills* opined that

the *EchoStar* court made such a ruling [regarding the temporal scope of the waiver] in the context of discovery from an outside attorney who had provided an opinion on the infringement issue after suit. There is no indication that the *EchoStar* court intended to extend this waiver to communication of trial counsel or to work product of trial counsel.<sup>40</sup>

Despite ruling that the waiver did not extend past the filing of the lawsuit and therefore did not extend to trial counsel, the court in *Indiana Mills* recognized that extending the scope of the waiver past the filing of the lawsuit and to communications with trial counsel is “an argument [that] could have merit” where “there is a fear that post-filing non-infringement opinions, whether they be of trial counsel, in-house counsel, or third-party outside counsel, differ from those pre-filing and have bearing on the reasonableness of the alleged infringer’s reliance on advice of counsel.”<sup>41</sup>

In contrast, other district courts have held that the waiver of privilege and work product immunity does extend to trial counsel. In *Informatica Corp. v. Business Objects Data Integration, Inc.*, a magistrate judge entered a discovery order holding that “[s]uch waiver applies to opinion counsel and trial counsel alike.”<sup>42</sup> The defendant elected to rely on the advice of counsel in combating a charge of willful infringement, and plaintiff sought discovery from both opinion counsel and trial counsel.<sup>43</sup> In *Informatica*, one of the attorneys who rendered an opinion on infringement was a member of the same firm representing the defendant as trial counsel.<sup>44</sup> The district court found that “according to the analysis in *EchoStar*, what is significant is the state of mind of [the defendant] and not the affiliation of [the defendant’s] attorneys, and that privilege has been waived with respect to pertinent communications and work product of all counsel in this case.”<sup>45</sup> Therefore, at least according to this *dicta*, it was immaterial whether trial counsel and opinion counsel were members of the same firm or different firms - the privilege had been waived with respect to both trial counsel and opinion counsel.

Still other district courts have sought middle ground. In *Intex Recreation Corp. v. Team Worldwide Corp.*, the district court took a reasoned approach. Trial counsel and opinion counsel were the same attorney, and the opinion had been given after the complaint had been filed.<sup>46</sup> The accused

<sup>38</sup> *In re Seagate Tech. LLC*, Misc. Docket No. 830, 2007 U.S. App. LEXIS 2457, at \*1 (Fed. Cir. Jan. 26, 2007). Oral argument in that case is currently scheduled for June 7, 2007. The Federal Circuit also invited the parties to brief the question of whether the Federal Circuit should reconsider the decision in *Underwater Devices, Inc. v. Morrison-Knudson Co.*, 717 F.2d 1380 (Fed. Cir. 1983) and the duty of care standard itself. *Id.*

<sup>39</sup> *Ind. Mills & Mfg., Inc. v. Dorel Indus., Inc.*, No. 1:04-cv-01102-LJM-WTL, 2006 U.S. Dist. LEXIS 34023, at \*\*19-20 (S.D. Ind. May 26, 2006) (opinion later withdrawn); see also *Ampex Corp. v. Eastman Kodak Co.*, No. 04-1373-KAJ, 2006 U.S. Dist. LEXIS 48702, at \*11 (D. Del. July 17, 2006) (“*EchoStar* did not even address the issue of communications with trial counsel.”).

<sup>40</sup> *Ind. Mills*, 2006 U.S. Dist. LEXIS 34023, at \*\*19-20.

<sup>41</sup> *Id.* at \*20, n.2.

<sup>42</sup> *Informatica Corp. v. Business Objects Data Integration, Inc.*, 454 F. Supp. 2d 957, 959 (N.D. Cal. 2006) (Larson, Mag. J.); see also *Celerity, Inc. v. Ultra Clean Holdings, Inc.*, 476 F. Supp. 2d 1159, 1164 (N.D. Cal. 2007) (where opinion counsel and trial counsel were attorneys from the same firm who were walled off from each other, the waiver extended to trial counsel as well as opinion counsel); *Iridex Corp. v. Synergetics, Inc.*, No. 4:05CV1916 CDR, 2007 U.S. Dist. LEXIS 7747, at \*\*2-4 (E.D. Mo. Feb. 2, 2007) (waiver applies to trial counsel and defendant must provide discovery regarding opinions of counsel communicated to defendant by trial counsel, except for those given solely to evaluate settlement options); *Computer Assocs. Int'l, Inc. v. Simple.com, Inc.*, No. 02 Civ. 2748 (DRH) (MLO), 2006 U.S. Dist. LEXIS 77077, at \*\*12-14 (E.D.N.Y. Oct. 20, 2006) (waiver applies to trial counsel, although not to “litigation strategy” as opposed to “advice of counsel”); *Beck Sys., Inc. v. ManageSoft Corp.*, No. 05 C 2036, 2006 U.S. Dist. LEXIS 53963, at \*16 (N.D. Ill. July 14, 2006) (The district court, although considering the opinion in *Indiana Mills*, concluded that *EchoStar* did not disturb the district court’s earlier decisions holding that the waiver extends to trial counsel, at least in the case of ongoing infringement.).

<sup>43</sup> *Id.* at \*6.

<sup>44</sup> *Id.*

<sup>45</sup> *Id.* at \*\*20-21.

<sup>46</sup> *Intex Recreation Corp. v. Team Worldwide Corp.*, 439 F. Supp. 2d 46, 52 (D.D.C. 2006).

infringer argued that if the waiver extended to all communications with trial counsel at all times, the patentee would be able to “inquire into counsel’s trial strategies, mental impressions and analysis, thereby piercing the zone of privacy for strategic litigation planning and allowing [the patent owner] to piggyback on [the accused infringer’s trial] preparations.”<sup>47</sup> While recognizing the language in *EchoStar* regarding temporal scope of the waiver, the district court found that “the Federal Circuit’s limited discussion of the temporal scope of waiver in footnote 4 of its *EchoStar* opinion is not, as [patentee] suggests, dispositive of this dispute.”<sup>48</sup> The district court held that the accused infringer would have to produce only testimony and documents from trial counsel that “question or contradict in any way the competence or validity of the opinions rendered.”<sup>49</sup>

In *Genentech, Inc. v. Insmid Inc.*, the district court ruled that the waiver of privilege extended to trial counsel, but only to “communications that are central and material to Defendants’ decision to engage in allegedly infringing activity.”<sup>50</sup> In a carefully-reasoned opinion, the magistrate judge considered both pre- and post-*EchoStar* cases and concluded that “although *EchoStar* does not expressly overrule those cases that decline to extend waiver to trial counsel, the thrust of the case is that a blanket immunity for trial counsel is not appropriate . . . particular solicitude [] must be paid to the policies that favor broad waiver.”<sup>51</sup> The “importance of protecting attorney work product and communication between attorney and client in the heat of litigation should not be lightly disregarded,” and therefore “a more refined balance is warranted.”<sup>52</sup> After concluding that the approach in *Intex* presents practical problems, the *Genentech* court attempts to draw a distinction between “significant opinions” and “lower level documents and communications” of trial counsel.<sup>53</sup> It remains to be seen whether either “middle ground” approach is workable in the context of a typical patent litigation.

## 2. Does Waiver Extend to All Defenses?

At least one district court has held that the waiver of privilege extends to all defenses, including infringement, invalidity, and unenforceability, regardless of the subject matter of the opinion. In *Intex*, the district court held that the scope of the waiver extended to any “document or opinion that embodies or discusses a communication . . . concerning whether [the patent-in-suit] is valid, enforceable, and infringed.”<sup>54</sup>

Other district courts have held that the waiver of privilege is limited to the subject matter of the opinion. In *Ampex Corp. v. Eastman Kodak Co.*, the district court denied a motion to compel all communications between defendant and its trial counsel bearing on the subject of infringement.<sup>55</sup> The district court carefully considered the language used in *EchoStar*. Specifically, the district court in *Ampex* explained that where the Federal Circuit stated that privilege was waived for “such communications regarding the same subject matter,” the

use of “such” leads one back to earlier language in that portion of the opinion, in which the court is emphasizing the unfairness of allowing a party to hold back an attorney’s opinion that is inconsistent with a different opinion it chooses to show the world. The modifier “such” thus strongly implies that the type of communications being discussed are opinions expressed in a manner comparable to the opinion that is disclosed, as was apparently the case in *EchoStar* itself.<sup>56</sup>

47 *Id.* at 51 (quotations omitted).

48 *Id.* at 52.

49 *Id.*

50 *Genentech, Inc. v. Insmid Inc.*, 442 F. Supp. 2d 838, 847 (N.D. Cal. 2006).

51 *Id.* at 846.

52 *Id.*

53 *Id.* at 847. Although the *Genentech* court had concerns regarding the practical application of the *Intex* standard, the *Genentech* approach may have similar flaws.

54 *Intex*, 439 F. Supp. 2d at 50-51; see also *Kimberly-Clark Corp. v. Tjco Healthcare Retail Group*, No. 05-C-985, 2007 U.S. Dist. LEXIS 5974, at \*\*4-5 (E.D. Wis. Jan. 26, 2007) (where privilege is waived as to a non-infringement opinion, the waiver extends to any invalidity or unenforceability opinion as well); *Affinion Net Patents, Inc. v. Maritz, Inc.*, 440 F. Supp. 2d 354, 356 (D. Del. 2006) (“Defendant has asserted advice of counsel as a defense to Plaintiff’s claims of willful infringement, and therefore, Defendant has waived the attorney-client privilege as to all communications relating to non-infringement, invalidity, and unenforceability.”).

55 *Ampex Corp. v. Eastman Kodak Co.*, No. 04-1373-KAJ, 2006 U.S. Dist. LEXIS 48702, at \*2 (D. Del. July 17, 2006); see also *CCC Info. Servs., Inc. v. Mitchell Int’l, Inc.*, No. 03 C 2695, 2006 U.S. Dist. LEXIS 87255, at \*\*17-18 (N.D. Ill. Dec. 1, 2006) (where privilege was waived as to an opinion on non-infringement, an opinion regarding damages contained in that opinion is not part of the same subject matter, i.e., liability, and is not within the scope of the waiver); *Autobytel, Inc. v. Dealix Corp.*, 455 F. Supp. 2d 569, 575 (E.D. Tex. 2006) (where an advice-of-counsel defense of non-infringement only is asserted, the waiver does not extend to defenses of invalidity and unenforceability).

56 *Ampex*, 2006 U.S. Dist. LEXIS 48702, at \*\*8-9 (citation omitted).

The *Ampex* court further reasoned that plaintiff's reading of *EchoStar*, extending the waiver to all communications bearing on the subject of infringement, would "demolish[] the practical significance of the attorney-client privilege, a result obviously at odds with other comments in *EchoStar*."<sup>57</sup>

Similarly, in *Informatica*, the magistrate judge held, without discussion, that the defendant "waives privilege for communications with counsel on the subject of the opinion or advice on which it relies."<sup>58</sup> This language would appear to reflect a narrower scope than all communications having any bearing on a defense to infringement. In affirming the magistrate's discovery order, the district court judge emphasized that the waiver did not extend to "trial strategy concerning validity, infringement, and enforcement in ways that do not implicate the advice-of-counsel defense" and the defendant need only disclose "opinions or advice that address the subject matter of the opinion."<sup>59</sup>

While the Federal Circuit in *EchoStar* attempted to provide guidance on the scope of the waiver of attorney-client privilege and attorney work product immunity, in just the few months since that decision, it appears any guidance was minimal at best.

### C. Loss of Effective Representation

The quandary facing defendants after *Knorr-Bremse* and *EchoStar* is so fraught with uncertainty that it could be argued to essentially deprive the defendant of effective representation. Defendants must choose whether to waive privilege with virtually no information regarding any potential liability determination and less than definitive information regarding the scope of any waiver. The risk that all communications with trial counsel could fall within the scope of a waiver raises the possibility that defendants will simply stop communicating with trial counsel, an undesirable result from a public policy standpoint.

Defendants are often forced to choose whether to waive privilege before a determination of liability. Thus, a defendant may choose to waive privilege in order to defend against a willfulness charge related to a patent that will ultimately be found not infringed or invalid. This "wasted" waiver helps no one – if nothing else, both parties have expended resources to conduct discovery regarding the advice of counsel. In the opposite situation, a defendant may choose not to waive, believing it has a strong case on liability. If the defendant then loses on liability, the defendant risks a willfulness finding as well, even if the defendant has obtained a competent opinion of counsel upon which it reasonably relied.

Where the defendant is forced to make a waiver decision without the benefit of a liability determination, and where, as discussed above, the scope of the waiver arguably may be so broad as to encompass nearly all communications between the defendant and trial counsel, the defendant and trial counsel may avoid communicating. Willfulness jurisprudence has been criticized for providing an incentive for companies to remain ignorant of competitors' patents.<sup>60</sup> Similarly, the Federal Circuit's opinions in *Knorr-Bremse* and *EchoStar* may provide an incentive for a defendant to avoid communicating with its attorneys. If attorney and client cannot engage in "full and frank communication,"<sup>61</sup> then the client is deprived of effective representation.

Allowing a party effective representation, particularly in a complex case such as a patent case, is good public policy:

Fully informed lawyers participating in the legal system as officers of the court sharpen the adversary process, thus improving the quality of judicial decision-making and the development of the law. By encouraging individuals to consult lawyers and disclose to them candidly and fully, the attorney-client privilege also

<sup>57</sup> *Id.* at \*10.

<sup>58</sup> *Informatica Corp. v. Bus. Objects Data Integration, Inc.*, 454 F. Supp. 2d 957, 965 (N.D. Cal. 2006) (Larson, Mag. J.).

<sup>59</sup> *Informatica Corp. v. Bus. Objects Data Integration, Inc.*, No. C 02-3378 JSW, 2006 U.S. Dist. LEXIS 58976, at \*8 (N.D. Cal. Aug. 9, 2006) (White, J.).

<sup>60</sup> See, e.g., Harold A. Borland, *The Affirmative Duty to Exercise Due Care in Willful Patent Infringement: We Still Want It*, 6 HOUS. BUS. & TAX L.J. 176, 186 (2005); Lemley & Tangri, *supra* note 37, at 1100; see also *Knorr-Bremse Systeme Fuer Nutzfahrzeuge GmbH v. Dana Corp.*, 383 F.3d 1337, 1351 (Fed. Cir. 2004) (en banc) (Dyk, J., dissenting-in-part).

<sup>61</sup> *Knorr-Bremse*, 383 F.3d at 1344 (citing *Upjohn Co. v. United States*, 449 U.S. 383, 389 (1981)).

allows the nation's legal profession to help individuals understand their legal obligations and facilitate their voluntary compliance with them. Such voluntary compliance is particularly important to a free society which neither has nor should want sufficient law enforcement resources to search out and punish every violation of every law.<sup>62</sup>

Patent cases are generally complex, and attorneys are particularly valuable in assisting their clients in understanding their obligations under the law.

In addition, encouraging full and free communication between attorney and client makes settlement more likely, since the parties will have a more accurate understanding of the merits of their relative positions and how those positions fit into the legal framework. The current state of Federal Circuit jurisprudence, however, may erode a defendant's ability to obtain effective representation and its corresponding societal advantages.

## II. CONGRESS' PROPOSAL

The flaws in willfulness litigation have not gone unnoticed by Congress. Two currently-identical bills pending in the House and the Senate, known as the Patent Reform Act of 2007, address many of these flaws.<sup>63</sup> Although neither bill may become law during the current 110th Congress, the fact that both bills' willfulness provisions are similar to those in earlier proposed patent legislation is an encouraging sign that a comparable provision may eventually be enacted as law.<sup>64</sup>

The Patent Act of 2005, pending in the House of Representatives during the 109th Congress, would have amended the law relating to willfulness in several respects.<sup>65</sup> First, the Patent Act of 2005 would have altered the affirmative duty of due care, and instead set forth certain actions taken by an infringer that may result in a finding of willfulness.<sup>66</sup> Second, it would have codified the decision in *Knorr-Bremse* abolishing the adverse inference where a defendant does not present evidence of advice of counsel.<sup>67</sup> Third, the Patent Act of 2005 would have prevented a plaintiff from even pleading willfulness until after a determination of liability: "A patentee may not plead, and a court may not determine, that an infringer has willfully infringed a patent before the date on which a determination has been made that the patent in suit is not invalid, is enforceable, and has been infringed by the infringer."<sup>68</sup>

The Patent Reform Act of 2006, pending in the Senate during the 109th Congress and introduced by Senators Hatch and Leahy, would have amended the law relating to willfulness in much the same way as the Patent Act of 2005.<sup>69</sup> The Patent Reform Act of 2006 would have also altered the affirmative duty of due care, and set forth the same actions by an infringer that may result in a finding of willfulness as does the Patent Act of 2005.<sup>70</sup> It also codified the decision in *Knorr-Bremse* abolishing the adverse inference where a defendant does not present evidence of advice of counsel.<sup>71</sup> Finally, the Patent Reform Act of 2006, like the Patent Act of 2005, would have prevented a plaintiff from pleading willfulness until after a determination of liability.<sup>72</sup>

In one important respect, the Patent Reform Act of 2006 differed from the Patent Act of 2005. The Patent Reform Act of 2006 would have made a willfulness determination a question to be determined by the court: "The court's determination of an infringer's willfulness shall be made without a jury."<sup>73</sup>

62 *In re Sealed Case*, 124 F.3d 230, 237 (D.C. Cir. 1997) (Tatel, J., dissenting); see also *Upjohn*, 449 U.S. at 389.

63 Neither proposed law addresses the scope of the waiver of attorney-client privilege and attorney work product should a defendant choose to rely on the advice of counsel.

64 See S. 3818, 109th Cong. Section 5(b) (2006); H.R. 2795, 109th Cong. Section 6(b) (2005).

65 H.R. 2795, 109th Cong. Section 6(b) (2005).

66 *Id.* at Section 6(b)(2).

67 *Id.* at Section 6(b)(3)(C).

68 *Id.* at Section 6(b)(4).

69 See S. 3818, 109th Cong. Section 5(b) (2006).

70 See *id.* at Sections 5(b)(2), 5(b)(3).

71 *Id.* at Section 5(b)(4)(C).

72 *Id.* at Section 5(b)(5).

73 *Id.*



The Patent Reform Act of 2007, pending in the Senate and the House of Representatives, is similar to both the Patent Act of 2005 and the Patent Reform Act of 2006.<sup>74</sup> First, the Patent Reform Act of 2007 alters the affirmative duty of due care, and instead sets forth certain actions taken by an infringer that may result in a finding of willfulness.<sup>75</sup> Second, it codifies the decision in *Knorr-Bremse* abolishing the adverse inference where a defendant does not present evidence of advice of counsel.<sup>76</sup> Third, the Patent Reform Act of 2007 would prevent a plaintiff from even pleading willfulness until after a determination of liability: "Before the date on which a court determines that the patent in suit is not invalid, is enforceable, and has been infringed by the infringer, a patentee may not plead and a court may not determine that an infringer has willfully infringed a patent."<sup>77</sup> Finally, the Patent Reform Act of 2007 would also make a willfulness determination a question for the court: "The court's determination of an infringer's willfulness shall be made without a jury."<sup>78</sup>

Congress may not enact the Patent Reform Act of 2007 during this session, or in its current form. The pending bill is a strong sign, however, that any future revision to the patent statute will include changes to willfulness law similar to those in the Patent Act of 2005, the Patent Reform Act of 2006, and the Patent Reform Act of 2007. Congress appears to recognize the flaws with willfulness law, and has proposed legislation that addresses those flaws.

### III. PROPOSED SOLUTIONS

It is unclear when, if ever, Congress will enact revisions to the patent statute, and whether the Federal Circuit's pending decision in *In re Seagate* will settle the numerous questions surrounding willfulness. There are, however, at least two ways that courts (both district courts and the Federal Circuit) can solve many of the problems presented by willfulness law after the Federal Circuit decisions in *Knorr-Bremse* and *EchoStar*. First, district courts could bifurcate trial on the willfulness issue and delay the defendant's decision on waiver and discovery on the issue of the advice of counsel until after a liability determination. Second, the Federal Circuit could reconsider its precedent and, as suggested in the Patent Reform Act of 2007, make willfulness, like inequitable conduct, a question for the court. While these solutions are not revolutionary,<sup>79</sup> they are particularly appropriate in light of the Federal Circuit's recent decisions and the proposed legislation. This paper adds to the growing chorus of dissatisfaction with willfulness law.

#### A. Bifurcation

District courts should bifurcate trial on the issue of willfulness and allow defendants to decide whether to rely on the advice of counsel after a determination of liability. Rule 42(b) of the Federal Rules of Civil Procedure authorizes the district court to order a separate trial of any claim or issue "in furtherance of convenience or to avoid prejudice, or when separate trials will be conducive to expedition and economy."<sup>80</sup> The district could, but need not, bifurcate discovery regarding willfulness factors other than the advice of counsel.

The Federal Circuit has favorably commented on such a bifurcation approach. In *Quantum Corp. v. Tandon Corp.*, the Federal Circuit recognized the difficult choice facing a defendant charged with willfulness and suggested bifurcation of the issue of willfulness as a potential solution:

Trial courts thus should give serious consideration to a separate trial on willfulness whenever the particular attorney-client communications, once inspected by the court *in camera*, reveal that the defendant is indeed confronted with this dilemma. While our court has recognized that refusal of a separate trial will not require

<sup>74</sup> See S. 1145, 110th Cong. Section 5(b) (2007); H.R. 1908, 110th Cong. Section 5(b) (2007).

<sup>75</sup> S. 1145, 110th Cong. Section 5(b)(2) (2007); H.R. 1908, 110th Cong. Section 5(b)(2) (2007).

<sup>76</sup> S. 1145, 110th Cong. Section 5(b)(3)(C) (2007); H.R. 1908, 110th Cong. Section 5(b)(3)(C) (2007).

<sup>77</sup> S. 1145, 110th Cong. Section 5(b)(4) (2007); H.R. 1908, 110th Cong. Section 5(b)(4) (2007).

<sup>78</sup> S. 1145, 110th Cong. Section 5(b)(4) (2007); H.R. 1908, 110th Cong. Section 5(b)(4) (2007).

<sup>79</sup> See, e.g., Taylor, *supra* note 37, at 346-48 (discussing bifurcation as a potential solution to the traditional defendant's dilemma); Lee & Cogswell, *supra* note 37, at 456-57 (same); John B. Pegram, *The Willful Patent Infringement Dilemma and the 7th Amendment*, 86 J. PAT. & TRADEMARK OFF. SOC'Y 271 (2004) (suggesting that willfulness should be a question for the court); Ira V. Heffan, *Willful Patent Infringement*, 7 FED. CIR. B.J. 115, 150-53 (1997) (suggesting both bifurcation and that willfulness should be a question for the court); Allan M. Littman, *The Jury's Role in Determining Key Issues in Patent Cases: Markman, Hilton Davis and Beyond*, 37 IDEA 207 (1997) (suggesting willfulness should be determined by the court).

<sup>80</sup> FED. R. CIV. P. 42(b).

reversal in every case involving attorney client communications bearing on willfulness, we have suggested the advisability of separate trials in appropriate cases.<sup>81</sup>

A number of district courts have already followed the approach laid out in dicta by the Federal Circuit in *Quantum*, including the Central District of California,<sup>82</sup> the Northern District of Illinois,<sup>83</sup> the Eastern District of Michigan,<sup>84</sup> District of New Jersey,<sup>85</sup> and the Eastern District of Pennsylvania.<sup>86</sup> Despite favorable commentary by the Federal Circuit and adoption by some district courts, this procedure is not yet widespread.

This proposal has the significant advantage that in many cases, the defendant will never need to channel resources into making the complex decision on whether to rely on the advice of counsel defense and waive privilege. In a substantial percentage of infringement suits, the defendant is found to have no liability, either because the patent(s)-in-suit are invalid or not infringed. The district court in *Pfizer Inc. v. Novopharm Ltd.* noted that “[b]ifurcation could be highly efficient, in that a verdict of no liability for infringement would render discovery and trial on the willfulness issue unnecessary.”<sup>87</sup> Although there might be some inefficiencies involved in either bringing back the same jury to hear the willfulness case or in impaneling a second jury, those inefficiencies would likely be outweighed by the number of times the willfulness case could be avoided altogether.

Additionally, this proposal has the advantage that attorney-client opinion evidence related to willfulness will not be able to influence the liability determination. Although impossible to measure, it is easy to imagine that a jury’s liability determination could be colored by exposure to evidence relating to a defendant’s first knowledge of a patent and subsequent failure to obtain an opinion.

Bifurcating willfulness for the purposes of trial and postponing the defendant’s decision until a determination of liability is a procedural step that district courts can take now. This step can be taken without further input from the Federal Circuit or Congress, and addresses many of the flaws with willfulness law.

## B. Willfulness Should be an Issue for the Court

As suggested in the Patent Reform Act of 2007, discussed above, willfulness should be a question for the courts. Although the law is that “[w]illfulness of behavior is a classical jury question of intent,”<sup>88</sup> there are several good reasons for the Federal Circuit to reconsider this precedent.<sup>89</sup>

First, the Federal Circuit should reconsider its position because the court and not the jury already determines whether infringement is willful for practical purposes. The consequence of a finding of willful infringement is the possibility of financial penalty via enhanced damages and the imposition of attorneys’ fees.<sup>90</sup> Yet it is the court, not a jury, who decides whether damages will be enhanced and whether or not attorneys’ fees will be awarded.<sup>91</sup> Moreover, when determining whether damages will be enhanced, the district court considers the same *Read* factors that a jury considers in determining willfulness.<sup>92</sup>

81 See *Quantum Corp. v. Tandon Corp.*, 940 F.2d 642, 644 (Fed. Cir. 1991).

82 See *Yamaha Hatsudoki Kabushiki Kaisha v. Bombardier Inc.*, 59 U.S.P.Q.2d (BNA) 1088 (C.D. Cal. 2001).

83 See *Pfizer Inc. v. Novopharm Ltd.*, 57 U.S.P.Q.2d (BNA) 1442 (N.D. Ill. 2000) (granting motion for bifurcation for reasons of efficiency).

84 See *Patent Holding Co. v. TG (USA) Corp.*, 46 U.S.P.Q.2d (BNA) 1566 (E.D. Mich. 1998) (recognizing potential for Quantum dilemma and deciding to review documents *in camera* before final determination).

85 See *Princeton Biochemicals, Inc. v. Beckman Instruments, Inc.*, 180 F.R.D. 254 (D.N.J. 1997).

86 See *B. Braun Med., Inc. v. Abbott Labs.*, 32 U.S.P.Q.2d (BNA) 1211 (E.D. Pa. 1994).

87 *Pfizer*, 57 U.S.P.Q.2d (BNA) at 1444.

88 *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1250 (Fed. Cir. 1989) (superseded in other part by California statute); see also *Biotec Biologische Naturverpackungen GmbH & Co. v. Biocorp, Inc.*, 249 F.3d 1341, 1356 (Fed. Cir. 2001); *SRI Int'l, Inc. v. Adv. Tech. Labs., Inc.*, 127 F.3d 1462, 1465 (Fed. Cir. 1997); *Natl Presto Indus., Inc. v. W. Bend Co.*, 76 F.3d 1185, 1192-93 (Fed. Cir. 1996); *Pall Corp. v. Micron Separations, Inc.*, 66 F.3d 1211, 1221 (Fed. Cir. 1995).

89 Commentators disagree as to whether making willfulness a question for the court would violate the right to a jury trial in the Seventh Amendment to the United States Constitution. See, e.g., DONALD S. CHISUM, CHISUM ON PATENTS Section 20.03[4][b][vii] (1999); Heffan, *supra* note 79, at 151-53; Pegram, *supra* note 79; Alan N. Herda, Comment, *Willful Patent Infringement and the Right to a Jury Trial*, 9 TEX. WESLEYAN L. REV. 181 (2003); 7 Louis L. Wu, Comment, *Enhanced Damages for Willful Patent Infringement—An Issue for Judge or Jury?*, 33 U.S.F. L. REV. 435 (1999). In any event, the Federal Circuit has thus far avoided the question of whether the Seventh Amendment includes a right to a jury trial on willfulness. See *Transmatic, Inc. v. Gulston Indus., Inc.*, 53 F.3d 1270, 1279 (Fed. Cir. 1995).

90 There may be some consequences to a defendant’s reputation from being found to have infringed willfully rather than merely infringed, but these consequences are neither certain nor ascertainable.

91 See 35 U.S.C. Sections 284, 285.

92 *Knorr-Bremse Systeme Fuer Nutzfahrzeuge GmbH v. Dana Corp.*, 383 F.3d 1337, 1342-43 (Fed. Cir. 2004) (en banc).

Because the court ultimately determines whether there will be any cost resulting from willful infringement, the court is already effectively determining whether infringement has been willful – if a defendant is found to infringe willfully but not subject to enhanced damages or attorneys' fees, the jury's determination is virtually meaningless. Numerous Federal Circuit decisions affirm district courts that, despite jury findings of willfulness, do not enhance damages by any amount.<sup>93</sup> Insisting that willfulness is a question for the jury rather than a question for the court when the court has the sole ability to impose any consequence (or no consequence at all) seems illogical. Reconsidering Federal Circuit precedent on this issue would merely bring jurisprudence in line with reality, an admirable goal whenever possible.

Second, the Federal Circuit should reconsider its position because making willfulness a question of equity for the court would alleviate the defendant's punitive dilemma regarding the advice of counsel defense and the waiver of attorney-client privilege. Because willfulness would not be tried to a jury, there would be few, if any, inefficiencies in allowing the defendant to wait until a liability determination to choose whether to rely on the advice of counsel and waive privilege. Limited discovery could then be conducted regarding the advice of counsel, and the willfulness issue tried to the court. In addition, the court would presumably be better able than a jury to refrain from drawing adverse inferences should no opinion be presented.

While continuing to maintain that willfulness is a question for the jury, the Federal Circuit recognizes the equitable nature of the willfulness analysis. In *SRI International, Inc. v. Advanced Technology Laboratories, Inc.*, the Federal Circuit, while ultimately holding that willfulness is a question of fact for the jury, also recognized that “the remedy of enhancement of damages not only serves its primary punitive/deterrent role, but in so doing it has the secondary benefit of quantifying the equities as between patentee and infringer.”<sup>94</sup> Similarly, in *National Presto Indus., Inc. v. West Bend Co.*, the Federal Circuit noted that “fairness as between patentee and infringer is a consideration in the determination of whether illegal behavior warrants an enhanced penalty,” but went on to insist that “the question does not thereby become ‘equitable.’”<sup>95</sup>

Courts, including the Federal Circuit, are also already experienced at analyzing equitable issues, including questions of intent, in patent cases. Inequitable conduct is an equitable question for the court.<sup>96</sup> A determination of inequitable conduct requires considering such fact issues as whether the patentee intended to deceive the United States Patent and Trademark Office.<sup>97</sup>

While reconsidering whether willfulness should be an issue of equity for the court sounds like a substantial change in the law, its transitional effects would be minimal since the court in effect already determines willfulness. In addition, trying the issue of willfulness to a court rather than a jury would allow a defendant to make a more informed decision regarding whether to rely on the advice of counsel defense.

#### IV. CONCLUSION

A defendant should not wield the attorney-client privilege as both a sword and shield. In the same vein, a patentee should not use a willfulness charge to deprive a defendant of effective representation, leaving the defendant both unarmed and defenseless.

Disentangling the “satellite litigation” of willfulness allegations from the meat of a patent suit, either through statute, bifurcation, or recognizing that willfulness should be an issue for the court to decide, addresses the continuing flaws in willfulness law in the wake of *Knorr-Bremse* and *EchoStar*.

<sup>93</sup> See, e.g., *Transclean Corp. v. Bridgewood Servs., Inc.*, 290 F.3d 1364 (Fed. Cir. 2002); *Electro Scientific Indus., Inc. v. Gen. Scanning, Inc.*, 247 F.3d 1341 (Fed. Cir. 2001); *Odetics, Inc. v. Storage Tech. Corp.*, 185 F.3d 1259 (Fed. Cir. 1999); *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448 (Fed. Cir. 1998) (en banc); *Laitram Corp. v. NEC Corp.*, 115 F.3d 947 (Fed. Cir. 1997); *Delta-X Corp. v. Baker Hughes Prod. Tools, Inc.*, 984 F.2d 410 (Fed. Cir. 1993); *Brooktree Corp. v. Adv. Micro Devices, Inc.*, 977 F.2d 1555 (Fed. Cir. 1992); *Modine Mfg. Co. v. Allen Group, Inc.*, 917 F.2d 538 (Fed. Cir. 1990).

<sup>94</sup> *SRI Int'l, Inc. v. Adv. Tech. Labs., Inc.*, 127 F.3d 1462, 1468 (Fed. Cir. 1997).

<sup>95</sup> *Nat'l Presto Indus., Inc. v. W. Bend Co.*, 76 F.3d 1185, 1193 (Fed. Cir. 1996).

<sup>96</sup> *Kingsdown Med. Consultants, Ltd. v. Hollister Inc.*, 863 F.2d 867, 876 (Fed. Cir. 1988) (with en banc “Resolution of Conflicting Precedent”) (citing *Garco Mfg. Inc. v. Herst Lighting Co.*, 820 F.2d 1209, 1212 (Fed. Cir. 1987)).

<sup>97</sup> See, e.g., *Molins PLC v. Tectron, Inc.*, 48 F.3d 1172, 1178 (Fed. Cir. 1995).

# UPDATES ON THE CORPORATE ATTORNEY-CLIENT PRIVILEGE

---

*David M. Brodsky\**  
*Latham & Watkins, LLP*  
*New York, NY*

## I. BACKGROUND

It has long been thought that the attorney-client privilege is one of the most sacrosanct and important evidentiary privileges available to lawyers and their clients. Because it encourages the free and open discussion between counsel and their clients, so as to allow full consideration of all legal issues that affect individuals and corporate entities alike. In recent years, however, some aspects of the attorney-client privilege have come under attack and increased scrutiny. Because one of the principal elements of the privilege is the ability of a client, corporate or individual, to refuse to reveal the nature of advice rendered by counsel, and the concomitant requirement that such counsel also refuse to reveal information communicated in order to seek advice or the advice so rendered, the privilege has long been viewed as hindering the ascertainment of the truth. But, in times past, this hindrance was thought necessary to encourage the full flowering of confidential communications between clients and their counsel.

This is no longer the case, especially in the case of the corporate attorney-client privilege. In recent years, there has been a concerted effort by federal, state, and local regulatory authorities to upend the traditional presumption that privileged information will remain confidential, with the result that a “culture of waiver” has set in, where virtually all corporations are being requested – or are expected – to waive or not assert the corporate attorney-client privilege to obtain the advantages associated with corporate cooperation with governmental investigations. In addition, because of inconsistent rulings in court proceedings regarding so-called “selective waiver,” disclosure by a corporation to an investigator or even to its external auditor are many times deemed to be general waivers as to all third parties, including counsel for adverse parties. The result is a substantial diminution of the protections of the corporate attorney-client privilege, and concomitant damage to the expectations of confidentiality which formerly accompanied consultations of corporate clients with their internal or external counsel.

This article will first analyze the corporate attorney-client privilege and work product doctrine (the “protections”) and the role of such protections in our society. Then, the article will deal with some of the significant developments adversely affecting the protections, and how some of the developments can be ameliorated so as to preserve the important principles underlying the corporate attorney-client privilege.

---

\* Mr. Brodsky is a member of Latham & Watkins LLP and co-chair of the firm's Global Securities Litigation and Professional Liability Practice Group. He is a member of the New York State Bar Association Task Force on the Attorney-Client Privilege and liaison to the American Bar Association Task Force on the Attorney-Client Privilege from the Consortium of Corporate Counsel.

## II. BRIEF SURVEY OF THE PROTECTIONS

### A. The Attorney-Client Privilege

The privilege is an evidentiary rule and protects confidential disclosures of information between lawyers and their clients so as to encourage full and frequent communications. Since the privilege generally thwarts full disclosure, attempts are frequently made to limit its scope through various exceptions. Therefore, it is vital to observe the requisites of the privilege so as to maximize the protections it provides. First, there must be a client, whether an individual or a corporation.<sup>1</sup> When a corporation is the client, corporate counsel must always distinguish between the individual acting as the corporation, and that same individual acting for herself. Careful advice must be given to such individuals to make certain that it is clear that corporate counsel represents only the corporate entity.

Second, there must be a lawyer who is rendering advice to the corporation and from whom or to whom confidential communications may flow. In this regard, whether counsel is inside or outside the corporation is irrelevant; what is relevant is the role that the attorney is playing. In the case of inside counsel, often there is a blurring between the attorney's role *qua* attorney, and such person's role *qua* business advisor. The latter role is not one where communications can be protected from disclosure. Obtaining legal, not business, advice as difficult as that may be to determine, must be the reason why corporate counsel was retained, and will be the basis for judging whether corporate counsel's communications will remain confidential.<sup>2</sup>

Third, it is the communications between attorney and client that are protected from disclosure, not the underlying facts that the attorney may uncover by reason of her work. Thus, corporate documents that convey factual information, upon which the corporate counsel then based her advice to the Company, are not protected from disclosure, although the advice would be.<sup>3</sup>

Fourth, the communication must be confidential in order to be protected. Presence of an unprotected third party in an otherwise confidential setting would constitute a waiver of the privilege's protection, as would communication to a third party of otherwise confidential information of a client.<sup>4</sup> Other persons assisting corporate counsel may also engage in communications that are protected from disclosure if it is clear that their communications are made for the purpose of obtaining legal advice from the lawyer.

Fifth, communication to and from similarly-situated individuals or corporate clients, or their or its counsel, may be protected by reason of the joint defense or common interest doctrine. The parties, in the first instance, must have entered into an agreement to cooperate among each other before such privilege will be extended.<sup>5</sup>

### B. The Work Product Doctrine

Protecting the confidentiality of work product likewise furthers vital public interests. "[T]he work product privilege [exists] . . . to promote the adversary system by safeguarding the fruits of an

1 *See Upjohn Co. v. U.S.*, 449 U.S. 383 (1981) (holding that privilege and work product protections may be asserted by corporations with respect to internal investigations); *Srougo v. BEA Assocs.*, 199 F.R.D. 515 (S.D.N.Y. 2001) ("The attorney-client privilege applies not only to individuals, but also to corporate entities.");

2 Courts have ruled that the communication by a client must be "made for the purpose of obtaining legal advice and directed to an attorney who has been consulted for that purpose." *Rosi v. Blue Cross & Blue Shields of Greater New York*, 73 N.Y.2d 588, 593 (N.Y. 1989) (internal quotations and citations omitted). And if the communication is from a lawyer to the client it must be "for the purpose of facilitating the rendition of legal advice or services, in the course of a professional relationship." *Id.*

3 *In re Grand Jury Subpoena Duces Tecum*, 731 F.2d 1032 (2d Cir. 1983) ("it is important to bear in mind that the attorney-client privilege protects communications rather than information; the privilege does not impede disclosure of information except to the extent that that disclosure would reveal confidential communications"); *ECDC Environment v. N.Y. Marine and General Ins. Co.*, No. 96 Civ. 6033, 1998 WL 614478 (S.D.N.Y. June 4, 1998) (finding that the documents were not subject to attorney-client privilege protection because they "merely convey factual information to plaintiff without any legal analysis or legal advice").

4 3 Weinstein's Federal Evidence Section 503.09 [1] ("Any disclosure of a lawyer-client communication to a third party for purposes inconsistent with maintaining confidentiality waives the attorney-client privilege as to that disclosure.")

5 *The Erosion of the Attorney-Client Privilege and the Work Product Doctrine in Federal Criminal Investigations*, Report of American College of Trial Lawyers, March 2002, at 15-16 (quoting *White-Collar Prosecutors Probe Joint Defense Agreements*, 1 The DOJ Alert 3, July 1991) ("DOJ Alert"); see also Report of the Am. Bar Ass'n Task Force on Attorney-Client Privilege (May 18, 2005), available at [www.abanet.org/buslaw/attorneyclient/materials/hod/report.pdf](http://www.abanet.org/buslaw/attorneyclient/materials/hod/report.pdf) ("ABA Task Force Report"). The Department of Justice and other law enforcement agencies have periodically attacked Joint Defense Agreement (JDAs) because prosecutors see that JDAs, "even unintentionally, [provide] an opportunity to get together and shape testimony." Prosecutors are cautioned, however, not to have a "knee-jerk reaction" against Joint Defense Agreements, unless there is a "specific reason to believe the agreement [is] being used for improper purposes." See DOJ Alert.

attorney's trial preparations from the discovery attempts of the opponent"<sup>6</sup> and "by affording an attorney 'a certain degree of privacy' so as to discourage 'unfairness' and 'sharp practices.'"<sup>7</sup> As one Supreme Court Justice wrote in a concurring opinion to the seminal decision supporting the doctrine, "[d]iscovery was hardly intended to enable a learned profession to perform its functions . . . on wits borrowed from the adversary."<sup>8</sup>

Protection of work product is codified in Federal Rule of Civil Procedure 26(b)(3), which extends protection to the work of a party's representatives, "including an attorney, consultant, surety, indemnitor, insurer, or agent" in anticipation of litigation or for trial. Work product is not discoverable by an opposing party absent a showing of "substantial need for the materials in the preparation of the party's case and [inability] without undue hardship to obtain the substantial equivalent of the materials by other means." But even when an opposing party makes this showing, courts must protect against disclosure of the "mental impressions, conclusions, opinions or legal theories of an attorney or other representative of a party."<sup>9</sup> It is *only when it would not be unfair* for an adversary to obtain that work product – *i.e.*, when the adversary meets its burden to show that it "has substantial need of the materials in the preparation of the party's case and is unable without undue hardship to obtain the substantial equivalent of the materials by other means"<sup>10</sup> – that the policy to protect work product will not apply.

Not just individuals but also companies expect that the work product of their counsel will be protected, and legitimately so. Increasingly, companies and (when the circumstances warrant) their audit committees or other independent committees use counsel to investigate evidence of alleged corporate or employee wrongdoing by interviewing company employees, identifying relevant documents, analyzing the facts and law and formulating conclusions and recommendations. Internal investigations, conducted by and at the direction of legal counsel, are a critical tool by which companies and their boards learn about violations of law, breaches of duty and other misconduct that may expose the company to liability and damages. Internal investigations are an essential predicate to enabling companies to take remedial action and to formulate defenses, where appropriate. Companies are therefore entitled to and afforded work product protection from adversaries, so long as the investigations are not merely being conducted in the ordinary course of business. One commentator has noted that the "general rationale for finding work product protection is that litigation is virtually assured if the investigation confirms the allegations. Since the corporation would be required to report the results to shareholders and government agencies, the possibility of a suit following is considered inevitable."<sup>11</sup>

Work product protection does not mean that, where internal investigations are conducted by counsel, all facts related to the issue under investigation are inherently protected against disclosure to auditors or third parties. The *facts*, including underlying documents, regarding an issue are properly discoverable, and routinely produced, in litigation. By contrast, what is protected from disclosure is the work performed, materials generated and considerations of the lawyers in connection with the investigation and any recommendations to the company. This is the heart of what is protected by the work product doctrine. The distinction among the facts, the law and the strengths and weaknesses of the company's position is an important one that is well-accepted in the law.<sup>12</sup>

6 *In re Raytheon Securities Litig.*, 218 F.R.D. 354, 359 (D. Mass. 2003) (quoting *United States v. Amer. Tel. & Tel. Co.*, 642 F.3d 1286, 1299 (D.C. Cir. 1980)).

7 Joint Drafting Committee of the American College of Trial Lawyers, *The Erosion of the Attorney-Client Privilege and Work Product Doctrine in Federal Criminal Investigations* (March 2002), at 6, quoting *Hickman v. Taylor*, 329 U.S. 495, 510-11 (1946).

8 *Hickman v. Taylor*, 329 U.S. 495, 516 (1946) (Jackson, J., concurring).

9 Fed. R. Civ. P. 26(b)(3).

10 Fed. R. Civ. P. 26(b)(3).

11 John William Gergacz, *Attorney-Corporate Client Privilege* Section 7.37 (West 2000), at 7-53 (reporting that "[m]ost of the cases hold that intracorporate investigations of possible corporate illegal activity are performed with sufficient anticipation of litigation to give rise to work product protection"). The author also reports that it is not only the inevitability of litigation, but also "the importance of not discouraging corporate self-investigation, [which] provides the underlying basis for the finding of work product protection." *Id.* at 7-54.

12 See *Sporck v. Peil*, 759 F.2d 312, 315 (3rd Cir. 1985) (lawyer's choice of documents with which to prepare deponent is work product even if the underlying documents themselves are not, "[b]ecause identification of documents as a group will reveal defense counsel's selection process, and thus his mental impressions...."); see also *In re Grand Jury Subpoenas Dated October 22, 1991 and November 1, 1991*, 959 F.2d 1158, 1166-67 (2d Cir. 1992) (noting that work product exception is only found when there is "real, rather than speculative concern that the thought process of [the client's] counsel... would be exposed"; and allowing production of all telephone records from a specified period) (internal citations and quotations omitted); *In re Grand Jury Subpoenas Dated March 19, 2002 and August 2, 2002*, 318 F.3d 379, 386-87 (2d Cir. 2003) (finding that lower court was correct in allowing discovery of disputed materials because producing party had failed to disclose any strategy ex parte to the district court judge, making it impossible for judge to determine whether the responsive subset of documents reflected lawyers' selection or was simply the product of document retention policies); *Shelton v. American Motors Corp.*, 805 F.2d 1323, 1326 (8th Cir. 1987) ("We hold that where, as here, the deponent is opposing counsel and has engaged in a selective process of compiling documents from among voluminous files in preparation for litigation, the mere acknowledgment of the existence of those documents would reveal counsel's mental impressions, which are work product.").

### III. THE PUBLIC INTEREST IN PRESERVING THE ATTORNEY-CLIENT PRIVILEGE AND WORK PRODUCT PROTECTION

It was formerly beyond serious discussion that the attorney-client privilege and work product doctrine as applied in the corporate context are “vital protections that serve society’s interests and protect clients’ Constitutional rights to counsel.”<sup>13</sup> A legal system that fails to assure business entities the benefits of the attorney-client privilege and work product protection denies those entities the effective assistance of counsel when potentially illegal corporate behavior is discovered within the organization. As the Supreme Court has stated, impairment of these privileges and protections would “not only make it difficult for corporate attorneys to formulate sound advice when their client is faced with a specific legal problem but also threaten to limit the valuable efforts of corporate counsel to ensure their client’s compliance with the law.”<sup>14</sup>

The strongest criticism of the attorney-client privilege – and, indeed, of any evidentiary privilege – had always been that, in court proceedings, potentially valuable evidence may be suppressed and the “truth” harder to find. “The social good derived from the proper performance of the functions of lawyers acting for their clients is believed to outweigh the harm that may come from the suppression of the evidence in specific cases.”<sup>15</sup> The Supreme Court long ago clarified that this social good extends to corporations as well as to individuals.<sup>16</sup>

Nonetheless, the upsurge of investigations into alleged corporate criminality has reignited the debate over the value of the privilege and the ability to have confidential communications between attorneys and clients. The risk of waiver by companies has increased in proportion to the volume of requests for disclosure of privileged materials. But it is precisely those confidential communications between corporate attorneys and the employees of the corporate client that are imperiled when the attorney-client privilege or work product doctrine is undermined. “Without reliable privilege protections, executives and other employees will be discouraged from asking difficult questions or seeking guidance regarding the most sensitive situations. Without meaningful privilege protections, lawyers are more likely to be excluded from operating in a preventive (rather than reactive) manner.”<sup>17</sup>

And it is not only corporate employees who will curtail – and have curtailed – the extent of their confidential communications with counsel to seek legal advice on business programs and strategies. It is our experience that company legal counsel (internal and outside) are curtailing their own activities, such as taking extensive notes at business meetings, for fear that if the subject of the business meetings were ever implicated in a governmental inquiry (where the company might not even be the “target”), such counsel’s notes would be turned over when the company waived the privilege and the counsel would be converted into a potential adverse witness against the company as client. Even outside counsel retained to conduct internal investigations have to be sensitive to procedures that might result in their becoming involuntary adverse witnesses. Those pressures create a potential conflict of interest between attorney and client that the privilege otherwise helps to prevent.

Protecting the confidentiality of work product likewise furthers vital public interests.<sup>18</sup> Work product protection supports a fair adversary system by “by affording an attorney ‘a certain degree of privacy’ so as to discourage ‘unfairness’ and ‘sharp practices.’”<sup>19</sup> The work-product doctrine is simply a recognition that a lawyer’s work on behalf of a client preparing a response to litigation or a potential claim – even when not subject to the attorney-client privilege – must also be protected, lest all lawyers be discouraged from conducting those preparations effectively, the clients be punished and their

13 See “The Decline of the Attorney-Client Privilege in the Corporate Context,” Survey Results, Presented to the United States Congress and the United States Sentencing Commission, March 2006.

14 *Upjohn Co. v. U.S.*, 449 U.S. 383, 392 (1981). This point was made forcefully in *Comments of the ABA’s Section of Antitrust Law On The Proposed Amendments To The Sentencing Guidelines For Organizations*, at 5-7, available at [http://www.abanet.org/antitrust/comments/2004/sentencing\\_guidelines0704.pdf](http://www.abanet.org/antitrust/comments/2004/sentencing_guidelines0704.pdf).

15 *United States v. United Shoe Mach. Corp.*, 89 F. Supp. 357, 358 (D. Mass. 1950); see *Trammel v. United States*, 445 U.S. 40, 50 (1980) (the privilege “promotes a public goal transcending the normally predominant principle of utilizing all rational means for ascertaining the truth”).

16 *Upjohn*, 449 U.S. at 389-90.

17 See “The Decline of the Attorney-Client Privilege in the Corporate Context,” Survey Results.

18 “[T]he work product privilege [exists] . . . to promote the adversary system by safeguarding the fruits of an attorney’s trial preparations from the discovery attempts of the opponent,” *In re Raytheon Securities Litig.*, 218 F.R.D. 354, 359 (D. Mass. 2003) (quoting *United States v. Amer. Tel. & Tel. Co.*, 642 F.3d 1286, 1299 (D.C. Cir. 1980)).

19 Joint Drafting Committee of the American College of Trial Lawyers, *The Erosion of the Attorney-Client Privilege and Work Product Doctrine in Federal Criminal Investigations* (March 2002), at 6, quoting *Hickman v. Taylor*, 329 U.S. 495, 510-11 (1946).

adversaries be unfairly rewarded. Those corporate clients (including their authorized representatives) who fear that the work product generated by their counsel in determining an appropriate response will be disclosed to their adversaries and promptly used against them will, not surprisingly, be reluctant to seek legal assistance at all much less provide information that will assist the attorney in providing such assistance.

#### IV. THE POLICIES OF THE UNITED STATES DEPARTMENT OF JUSTICE

Since the mid-1990s and continuing to date, the principal law enforcement and regulatory authorities in the United States have developed policies and guidelines that are designed to induce corporations and other business entities to waive, or not assert, applicable attorney-client and work-product privileges and protections.<sup>20</sup> First, in 1998, after several years of informal policies at various United States Attorney's Offices, principally the Southern District of New York, the Department of Justice formally adopted what came to be known as the "Holder Memorandum," after Eric Holder, then Deputy Attorney General of the United States. The Holder Memorandum, although advisory, set forth standards by which a corporation would be judged to be cooperative in a criminal investigation. One factor was whether the Corporation waived or did not assert privileges protecting the confidentiality of communication.

Then, in 2002, then Deputy-Attorney General, Larry Thompson, promulgated a revision of the Holder Memorandum, this time making mandatory the use of the factors in judging whether a corporation was sufficiently cooperative, including whether applicable principles were waived or not asserted. In 2004, following the general trend of policy reflected in the Thompson Memorandum, the United States Sentencing Commission adopted an amendment that a corporation's waiver of the attorney-client privilege and work product protections would be a prerequisite for obtaining a reduction by a corporation in its culpability score.

The adoption of these policies by the Department of Justice and other regulatory entities pushed aside historic policies protecting privilege and work-product in favor of policies promoting cooperation with governmental agencies and maximizing the effectiveness and efficiency of governmental investigations.<sup>21</sup> Companies formerly expected that the work product of their counsel prepared as a result of an internal investigation, and advice given as a result of such investigation, would be protected have come to learn that, upon the initiation of a governmental inquiry, whether formal or informal, whether the company is a target or not, such expectations of confidentiality were illusory. Internal investigations, conducted by and at the direction of legal counsel, that are a critical tool by which companies and their boards learn about violations of law, breaches of duty and other misconduct that may expose the company to liability and damages, and an essential predicate to enabling companies to take remedial action and to formulate defenses, where appropriate, no longer have clear and predictable protections of confidentiality in the current environment, viewed as a "culture of waiver."<sup>22</sup> At the height of the investigative storm, protected information is routinely expected to be made available to government authorities, sometimes, at the authorities' request, on a day-to-day basis during the internal investigations. Under such governmental policies, companies perceived they did not realistically have the option to preserve the confidentiality upon which an effective attorney-client relationship is so heavily dependent and otherwise protected by the privilege and doctrine, or they would run the considerable risk of being deemed "uncooperative" by the governmental authority – a characterization that is tantamount to a virtual corporate death sentence<sup>23</sup> or, at least, extraordinarily financially punitive.

20 See United States Attorneys' Criminal Resource Manual, Art. 162, Section VI.B; United States Sentencing Guidelines Manual Section 8C2.5(g)(2001); the SEC's Seaboard Report, available at <http://www.sec.gov/litigation/investreport/34-44969.htm>; see also the EPA Voluntary Disclosure Program, the HHS Provider Self-Disclosure Protocol, and the Department of Justice Antitrust Corporate Leniency Policy.

21 Committee Note to Proposed Amendments to the Federal Rules of Evidence (Rule 502), at 8; see also *In re Columbia/HCA Healthcare Corp. Billing Practices Litigation*, 293 F.3d 289, 314 (6th Cir. 2002) (Boggs, J., dissenting) ("the public interest in easing government investigations counsels against holding the attorney-work-product privilege waived when the holder of the privilege discloses privileged information to the government... a limited disclosure pursuant to a government agency's investigatory request ought not waive the privileges as to all other parties...").

22 See text at fn. 26.

23 As in the case of Arthur Andersen.



There are a variety of reasons why such authorities adopted such policies and, for a fuller discussion of them, we refer to the Report of the ABA's Task Force on the Attorney-Client Privilege,<sup>24</sup> and to the report by the Joint Drafting Committee of the American College of Trial Lawyers, *The Erosion of the Attorney-Client Privilege and Work Product Doctrine in Federal Criminal Investigations* (March 2002).<sup>25</sup> Regardless of the reasons proffered, the result at the Department of Justice, the Securities and Exchange Commission, the Commodity Futures Trading Commission, and other regulatory and self-regulatory agencies, as well as many state attorneys general offices and state regulatory agencies, has been a marked increase in the compelled, requested, suggested, or (pragmatically inevitable) "voluntary" waivers of the privilege and the work-product doctrine, in order to further enhance the likelihood that the company will avoid significant prosecution or regulatory action. The surge in such waivers has been well documented; a recent survey administered jointly by the Association of Corporate Counsel,<sup>26</sup> an organization representing nearly 19,000 public companies, and the National Association of Criminal Defense Lawyers<sup>27</sup> found that

- Nearly 75% of both inside and outside counsel state that, in their experience, government agencies expect a company under investigation to waive legal privileges (1 percent of in-house counsel and 2.5% of outside counsel disagreed with the statement);
- Of the respondents who confirmed that they or their clients had been subject to investigation in the past five years, approximately 30% of in-house counsel and 51% of outside counsel said that the government expected waiver in order to engage in bargaining or be eligible for more lenient treatment; and
- Of those who had been investigated, 55% of outside counsel said the privilege waiver was requested either directly or indirectly; 27% of in-house counsel confirmed that experience.

Of the over 675 responses to the survey, almost half of the general counsels responding on behalf of public and private companies have experienced some kind of privilege erosion, caused by the government's policies. Of these companies, by far the most were not from global companies with high visibility, but rather from a wide variety of differently-sized businesses. After more than a decade of increased pressure, explicit and implicit, on companies to waive the attorney-client privilege and work-product protections, there has emerged a "culture of waiver" in which government agencies expect a company under investigation to waive legal privileges, and many companies do so, most without even being asked any longer but knowing there is no practical alternative to doing so.

Recently, in response to rising criticism of the effects of these policies, some changes to these policies have begun to occur. First, in October 2004, the ABA established a Task Force on the Attorney-Client Privilege to examine the protections, "the circumstances in which competing objectives are currently being asserted by governmental agencies and others to override the privilege," and manners for achieving balance.<sup>28</sup> Following this, the Task Force held public hearings around the country, taking testimony from interested persons with familiarity with governmental policies and their effects, and preparing analyses of the relevant issues affecting the privilege.<sup>29</sup>

Then, in 2005 and early 2006, after hearings were held at which representatives of the American Bar Association, the Association of Corporate Counsel, and the U.S. Chamber of Commerce voiced objections to policies adopted in earlier years, the U.S. Sentencing Commission

<sup>24</sup> 60 Bus.Law. 1029 (May 2005); also available at <http://www.abanet.org/buslaw/attorneyclient/home.shtml>. The "Recommendation" approved by the ABA House of Delegates and outlined in footnote 4, *supra*, and not the related "Report" cited herein, constitutes official ABA policy.

<sup>25</sup> Available at <http://www.aatl.com>.

<sup>26</sup> "The Decline of the Attorney-Client Privilege in the Corporate Context." Survey Results, Presented to the United States Congress and the United States Sentencing Commission, March 2006, [http://www.nacdl.org/public.nsf/whitecollar/wcnews024/\\$FILE/A-C\\_PrivSurvey.pdf](http://www.nacdl.org/public.nsf/whitecollar/wcnews024/$FILE/A-C_PrivSurvey.pdf), and <http://www.acca.com/public/attyclntprvlg/coalitionsstestimony031506.pdf> ("Survey Results").

<sup>27</sup> The ABA and several other organizations provided active participation and access to their members in the survey process.

<sup>28</sup> See ABA Presidential Task Force on the Attorney-Client Privilege: Mission Statement, available at <http://www.abanet.org/buslaw/attorneyclient/home.shtml> (last visited August 29, 2006).

<sup>29</sup> See materials collected at <http://www.abanet.org/poladv/acprivilege.htm>

voted unanimously to reverse the 2004 amendment to the commentary for Section 8C2.5 of the Organization Sentencing Guidelines that had encouraged prosecutors to require companies and other organizations to waive their attorney-client privilege and work product protections as a condition for receiving credit for cooperation at sentencing. The Commission stated that such a deletion was made because it “could be misinterpreted to encourage waivers” by corporations.<sup>30</sup> The changes proposed were not modified by Congress and became effective on November 1, 2006.

The trend in criticism continued when, in August 2006, following the action by the U.S. Sentencing Commission, the ABA unanimously passed resolutions at its annual meeting, supporting the preservation of the attorney-client privilege and work product doctrine and opposing government policies and procedures that “have the effect of eroding the constitutional and other legal rights” of employees, past or present, if that employee decides to exercise his or her Fifth Amendment right against self-incrimination.” Arguments made in support of such resolutions included discussion of using impermissible factors to determine whether a corporation was cooperative with a governmental investigation, including whether the organization: (1) provides counsel to an employee; (2) participates in a joint defense and information sharing agreement with an employee or others with whom the organization has a common interest; (3) shares its own records or historical information about the conduct under investigation with an employee or other party involved in the investigation; or (4) declines to fire or otherwise sanction an employee who exercised his or her Fifth Amendment rights in response to government requests for information.

In the midst of this discussion, Federal District Judge Lewis Kaplan began to hold hearings in the KPMG Tax Shelter indictment, resulting in two well-publicized decisions regarding the over-breadth of the Thompson memorandum in its application and illegitimate use by the Department of Justice.<sup>31</sup>

At the same time, three former Attorneys General, three former Deputy AGs and four former Solicitors General urged that the DOJ revise its policies, specifically, the Thompson Memorandum, to state affirmatively that waiver of the attorney-client privilege and work product doctrine should not be factors in determining whether a corporation is being cooperative. The former DOJ officials call the “Thompson Memorandum” “seriously flawed” and argue that it “undermines, rather than enhances, compliance with the law.”<sup>32</sup>

## V. RESPONSE TO THE CRITICISM

In response to this intense criticism, there were two notable developments: first, Senator Arlen Specter (R-Pa.), then-chairman of the U.S. Senate Judiciary Committee, introduced legislation, entitled “Attorney-Client Privilege Protection Act of 2006,” to prevent the Department of Justice from demanding a waiver of the corporate attorney-client privilege;<sup>33</sup> and second, the Deputy Attorney General, Paul McNulty, issued a revision of the Thompson Memorandum.<sup>34</sup>

<sup>30</sup> <http://www.uscc.gov/2006guid/Finalamend2006.pdf>

<sup>31</sup> *In U.S. v. Stein*, 435 F.Supp.2d 330 (S.D.N.Y. 2006). Judge Kaplan dealt extensively with the issue of whether the Thompson Memorandum had had the effect, intended or otherwise, of forcing a large firm to withhold advancing of attorneys’ fees to former partners and employees, who were subjects and targets of a Grand Jury investigation, in order to earn for itself the label of having fully cooperated with the Government’s investigation, and thus be spared prosecution. The Court held that the government violated the Fifth and Sixth Amendment rights of the individual defendants by causing the firm—under threat of indictment and destruction of the firm—to depart from its uniform prior practice of paying the legal expenses of its personnel in all cases in which they were named in consequence of their activities on behalf of the firm. It found that the firm would have paid those expenses—whether legally obliged to do so or not—but for the government’s improper conduct.

In a later opinion in the same case, *U.S. v. Stein*, --- F.Supp.2d ---, 2006 WL 2060430, SDNY, 2006, July 25, 2006 the Court suppressed statements made by certain employees during the investigation, as having been coerced by the government in violation of their Fifth Amendment privilege against self-incrimination. The Court found that the government was responsible for the pressure put on the employees, by having threatened their employer’s existence, thus causing the firm to exert “substantial pressure on its employees to waive their constitutional rights.” The Court explicitly linked the firm’s pressure on their employees to the effects of the Thompson Memorandum: “In the words of its chief legal officer, [the firm] did everything it could to be able to say at the right time and with the right audience, we’re in full compliance with the Thompson Memorandum.”

<sup>32</sup> <http://online.wsj.com/article/SB115768279330257255.html>

<sup>33</sup> On January 4, 2007, Senator Specter re-introduced the legislation in the new 110th Congress and the “Attorney-Client Privilege Protection Act of 2007” was referred to the Senate Judiciary Committee for consideration. <http://thomas.loc.gov/cgi-bin/bdquery/z?d110:s:00186>

<sup>34</sup> See Principles of Federal Prosecution of Business Organizations, U.S. Deputy Attorney General Paul J. McNulty (December 12, 2006) (“McNulty Memorandum”), available at [http://www.usdoj.gov/dag/speech/2006/mcnulty\\_memo.pdf](http://www.usdoj.gov/dag/speech/2006/mcnulty_memo.pdf).

### A. The Specter Legislation

The proposed bill has the support of a diverse group, including the American Bar Association, National Association of Criminal Defense Attorneys, United States Chamber of Commerce, the American Civil Liberties Union, and the Association of Corporate Counsel. The bill would bar prosecutors from forcing companies to waive their attorney-client privilege over internal documents in order to avoid criminal charges, a key part of the Thompson Memo. In general, the proposed Act provides that an agent or attorney of the United States shall not

“demand, request, or condition treatment on the disclosure by an organization, or person affiliated with that organization, of any communication protected by the attorney-client privilege or any attorney work product...”

and

“condition a civil or criminal charging decision relating to an organization ...on, or use as a factor in determining whether an organization...is cooperating with the Government” any of the following:

- any valid assertion of the attorney-client privilege or privilege for attorney work product;
- the provision of counsel to an employee of that organization;
- the entry into a joint defense agreement with an employee of that organization, or
- failure to terminate any employee because of that employee’s decision to exercise his constitutional rights.

In addition, the bill would foreclose the Government from considering any of the foregoing actions when evaluating whether a corporation should be entitled to “cooperation credit.

The bill is responsive to the three specific issues raised by numerous witnesses before the U.S. Sentencing Commission Hearings, referred to above, namely agents of the Government should not demand or encourage an organization to refuse to indemnify an employee for her legal expenses, should not demand that the organization refuse to provide information to employees, and should not request or demand that an organization sanction an employee for invoking his Fifth Amendment privilege against self-incrimination. After having been introduced, the bill was sent to the Senate Judiciary Committee for further consideration.<sup>35</sup>

### B. The McNulty Memorandum

The “McNulty Memorandum” reaffirms many of the factors to be considered by federal prosecutors when conducting corporate investigations, considering whether to indict corporations and considering corporate plea agreements.<sup>36</sup> Of particular significance to any corporation, however, are the restrictions the new policy places on federal prosecutors with respect to: (1) requesting the disclosure of materials protected by the attorney-client privilege or work product doctrine, and considering whether the corporation makes such disclosure in deciding whether to indict the corporation; and (2) considering the corporation’s advancement of legal fees to its employees and other agents in deciding whether to indict the corporation.

<sup>35</sup> Senator Patrick Leahy, D-Vt., the new chair of the Judiciary Committee, has not yet announced his position with respect to the Specter bill. In a written statement following the release of the McNulty Memorandum, Senator Leahy made it clear that he welcomed the new prosecutorial standards and was “pleased that the Justice Department has heeded bipartisan criticism ... and moved away from its most excessive practices in corporate fraud investigations.” However, he also stated that he remained “concerned that, depending on how the new policies are implemented, prosecutors may still be able to inappropriately consider a corporation’s waiver of this important privilege” and that he would “continue to monitor the implementation of this new policy...” <http://www.acc.com/public/attvclientpriv/mcnultymemosenleahystatement.pdf>

<sup>36</sup> For a complete analysis of the McNulty Memorandum, see Shapiro and Malioneck, “The ‘Value of Cooperation’ - The Potential Impact of the McNulty Memorandum on DOJ Investigations,” <http://www.law.com/jsp/nylj/PubArticleNY.jsp?hubtype=outsideCounsel&id=1167127307792>

The new policy makes clear that most attorney-client communications “should only be sought in rare circumstances.”<sup>37</sup> It also establishes an approval process by which prosecutors seeking to request the disclosure of attorney-client protected information must establish a “legitimate need” for the information and obtain the approval of a supervisor, including in many cases the Deputy Attorney General himself. Importantly, if the corporation decides not to disclose protected information, other than purely factual information, “prosecutors will not view [the decision] negatively in making a charging decision.” Prosecutors may, however, give cooperation credit to corporations that decide to waive attorney-client protections.

In addition, the guidelines outlined by the McNulty Memorandum “now generally prohibit prosecutors from considering whether a corporation is advancing attorneys’ fees to employees or agents under investigation or indictment,” when making a charging decision. Only “in extremely rare cases,” when “the totality of the circumstances show” that fee advancement is being used “to impede a government investigation” may prosecutors consider this factor.<sup>38</sup>

The new policy also makes distinctions between different types of privileged materials (i.e., “Category I”, consisting mainly of factual work product materials and “Category II”, consisting of “attorney-client communications or non-factual attorney work product” materials) and establishes a separate approval standard and process for each type. The Memorandum also states that prosecutors cannot deny cooperation credit if a company refuses to waive but allows the government to grant credit to those companies that do waive, a provision that some commentators have labeled a “Catch-22.”<sup>39</sup>

Notwithstanding the changes from the Thompson Memorandum, the McNulty Memorandum still reserves for the Department of Justice the right to consider whether the organization was justified in not waiving the attorney-client privilege and attorney work-product, and makes it explicit that, while “waiver...is not a prerequisite to a finding that a company has cooperated...the disclosure of privileged information may be critical in enabling the government to evaluate the accuracy and completeness of the company’s voluntary disclosure.”<sup>40</sup> As one commentator has noted, “an examination of [the factors set forth in the McNulty memorandum] makes it clear that it would be easy for a prosecutor to determine that there is a legitimate need [for a waiver] in virtually any and every case.”<sup>41</sup>

## VI. THE ROLE OF AUDITORS

At the same time as companies were facing issues affecting their ability to rely on the attorney-client privilege because of governmental policies, they also began to deal with a different kind of pressure from their external auditors. Over the past decade, legal events and changes in business climate have altered perceptions regarding the role of independent auditors as “gatekeepers” of the financial markets. The pressures on independent auditors to require unfettered access to information possessed by their public company clients have increased through the fallout from corporate scandals, including legislation and increased regulation. With each increase in the pressures faced by auditors to demand information from their clients, the same clients have faced corresponding pressures to disclose privileged information to auditors in performance of their duties under generally acceptable auditing standards and under applicable laws.

We have been at similar crossroads before. In 1975, the legal and accounting professions reached an accord, often referred to as the “Treaty,” that was intended to balance seemingly competing interests: a client’s desire to maintain the protections of the attorney-client privilege and work product doctrine, on the one hand, and the auditor’s need for detailed corporate information related to the financial statements, on the other hand. The Treaty recognizes that maintaining the protections is

37 See McNulty Memorandum, *supra* n.33, at p. 10; see also Prepared Remarks of Deputy Attorney General Paul J. McNulty at the Lawyers for Civil Justice Membership Conference Regarding the Department’s Charging Guidelines in Corporate Fraud Prosecutions (Dec. 12, 2006), available at [http://www.usdoj.gov/speech/2006/dag\\_speech\\_061212.htm](http://www.usdoj.gov/speech/2006/dag_speech_061212.htm).

38 *Id.*

39 ABA Journal E-Report, “Thompson Memo Changes Not Enough, ABA Says,” December 15, 2006

40 See McNulty Memorandum, *supra* n.33, at p. 8.

41 N. Richard Janis, “The McNulty Memorandum: Much Ado About Nothing,” Washington Lawyer, February 2007, at 38.

entirely consistent with the function of auditors and their ability to obtain the information that they need to conduct proper audits.<sup>42</sup>

The mechanics by which the Treaty struck this balance involved limiting the range of acceptable disclosures that lawyers may make to auditors with the client's informed consent, and thus defining the scope of what the auditors may request from lawyers regarding confidential attorney information.<sup>43</sup> In 1977, the AICPA reaffirmed this protection and limitation regarding auditor access to privileged information and work product maintained by the client.<sup>44</sup>

More than just documenting these mechanics, however, the Treaty was an acknowledgement that there were common interests among companies, counsel and auditors.

*The public interest in protecting the confidentiality of lawyer-client communications is fundamental.* The American legal, political and economic systems depend heavily upon voluntary compliance with the law and upon ready access to a respected body of professionals able to interpret and advise on the law. The expanding complexity of our laws and governmental regulations increases the need for prompt, specific and unhampered lawyer-client communication. The benefits of such communication and early consultation underlie the strict statutory and ethical obligations of the lawyer to preserve the confidences and secrets of the client, as well as the long-recognized testimonial privilege for lawyer-client communication.<sup>45</sup>

Thus, while auditors require access to certain attorney-client information, they have also long recognized the importance of the privileges by cooperating in this "Treaty" designed to uphold the public interest in protecting these privileges.

While auditors historically have planned and performed their audits to obtain reasonable assurance that a company's financial statements are not materially misstated due to error or corporate fraud – and auditors continue to do so – recent developments in federal law and policy have strongly encouraged auditors to intensify their vigilance. The corporate scandals of 2001 and 2002 sparked legislative action by Congress, rule-making and enforcement initiatives by the SEC, standard-setting by the PCAOB, and initiatives by other oversight bodies, all of which have heightened the scrutiny over auditors' procedures to verify company positions and impacted generally accepted auditing standards ("GAAS") and how auditors apply GAAS.<sup>46</sup>

42 American Bar Association, "Statement of Policy Regarding Lawyers' Responses to Auditors' Requests for Information" (1975), available at <http://www.abanet.org/buslaw/catalog/5070426i/secure.html>.

Thoughtful law review articles discuss the tensions that led to the Treaty, including incidents of auditors asking lawyers open-ended questions seeking general information about the client's potential illegal acts and liability exposures. See Erbstoesser and Matson, *Lawyers' Letters to Auditors*, Ch. 8, Drafting Legal Opinion Letters, at 366, nn. 1 & 2 (2d ed. 1992); Deer, *Lawyers' Responses to Auditors' Requests for Information*, 28 Bus. Law. 947 (1973). The ABA Statement of Policy and SAS 12 ended these types of broad requests by clarifying that GAAS did not require them.

The Treaty involves three pieces of professional literature. The obligation of lawyers to limit their responses to auditor inquiries is set forth in the ABA Statement of Policy. The obligation of clients to accrue for and/or disclose loss contingencies properly is set forth in FAS 5, which is part of generally accepted accounting practices ("GAAP"). See Financial Accounting Standards Board, Statement of Accounting Standards No. 5: Accounting for Contingencies (March 1975). The obligation of auditors to inquire concerning litigation, claims and assessments is governed by GAAS and, specifically, SAS 12, adopted by Auditing Standards Executive Committee of the American Institute of Certified Public Accountants ("AICPA") in the wake of the ABA Statement of Policy. See AICPA, Auditing Standards Board, Statement on Auditing Standards No. 12: Inquiry of a Client's Lawyer Concerning Litigation, Claims and Assessments (Jan. 1976) (codified in AICPA Professional Standards, AU Section 337). The ABA Statement of Policy is an exhibit to SAS 12.

43 Pursuant to the Treaty, a lawyer may provide information to a client's auditors on matters to which the lawyer has devoted substantive attention regarding overtly threatened or pending litigation and, with the client's further specific consent, regarding unasserted possible claims or assessments or contractually-assumed obligations, and may provide specific confirmations regarding the lawyer's role for the client. Only in rare circumstances may the lawyer express to the auditors any professional judgment regarding the potential outcome of the matters. The lawyer may only provide information and evaluation of unasserted possible claims specifically identified by the client if the client has determined that it is "probable" the claims will be asserted, that there is a "reasonable possibility" that the outcome will be unfavorable and that the resulting liability will be material to the client's financial condition. ABA Statement of Policy, para. 5.

44 See AICPA Professional Standards, AU Section 9337 (4), Documents Subject to Lawyer-Client Privilege (March 1977). The interpretive release poses the question: "[SAS 12 states:] 'Examine documents in the client's possession concerning litigation, claims, and assessments, including correspondence and invoices from lawyers.' *Would this include a review of documents at the client's location considered by the lawyer and the client to be subject to the lawyer-client privilege?*" and answers as follows: "No. Although ordinarily an auditor would consider the inability to review information that could have a significant bearing on his audit as a scope restriction, in recognition of the public interest in protecting the confidentiality of lawyer-client communications, [SAS 12] is not intended to require an auditor to examine documents that the client identifies as subject to the lawyer-client privilege." (Emphasis added)

45 ABA Statement of Policy, Preamble (emphasis added).

46 Former SEC Enforcement Director Stephen M. Cutler referred to auditors in 2004 as one of the three principal "gatekeepers" in our capital markets, or "sentries of the marketplace." See Stephen M. Cutler, Director of the Division of Enforcement at the SEC, Remarks at the UCLA School of Law, Los Angeles, CA (September 20, 2004), "The Themes of Sarbanes-Oxley as Reflected in the Commission's Enforcement Program" (transcript available at <http://www.sec.gov/news/speech/spch092004smc.htm>).

In January 2002, then-SEC Chairman Harvey Pitt, discussing what he called the “Enron situation,” directed strong rhetoric towards auditors:

[T]here is a need for reform of the regulation of our accounting profession. We cannot afford a system, like the present one, that facilitates failure rather than success. Accounting firms have important public responsibilities. We have had too many financial and accounting failures. ... [T]he potential loss of confidence in our accounting firms and the audit process is a burden our capital markets cannot and should not bear.<sup>47</sup>

In 2004, Scott Taub, then-Deputy Chief Accountant and now Chief Accountant for the SEC, suggested that auditors should seek out privileged information in auditing reserves or accruals for litigation losses and tax contingencies under FAS 5:

The difficulty in auditing [loss contingency accruals under FAS 5], however, should cause the auditor to spend more time on them, not less. *If a company's outside counsel is unwilling or unable to provide its expert views, the auditor should consider whether sufficient alternate procedures can actually be performed to allow the audit to be completed.*<sup>48</sup>

In a report to Congress, mandated by Section 704 of the Sarbanes-Oxley Act, the SEC reported that, of 515 enforcement actions in total for the five years prior to 2003, 18 actions were filed against audit firms and 89 against individual auditors.<sup>49</sup> In the vast majority of these actions, auditors were sanctioned, in the SEC's words, for “failing to gain sufficient evidence to support the issuer's accounting, failing to exercise the appropriate level of skepticism in responding to red flags, and failing to maintain independence.”<sup>50</sup> The 704 Report concludes that “audit failures most often arise from auditors accepting management representations without verification, truncating analytical and substantive procedures, and failing to gain sufficient evidence to support the numbers in the financial statements.”<sup>51</sup>

The PCAOB has signaled an intention to be tough-minded in enforcing its mandate scrutinize auditors. In an August 2, 2004 interview, William McDonough, then Chairman of the PCAOB, stated his view on whether it is the auditor's *obligation* to detect client fraud.<sup>52</sup> He said:

We have a very clear view that it is their job [to detect fraud]. If we see fraud that wasn't detected and should have been, we will be very big on the tough and not so [big] on the love. ... [A]uditors [need to] understand that, with relatively few exceptions, they should find it. To me, the relatively few exceptions are those cases where you would have some extremely dedicated, capable crooks. In most cases, though, the crooks either are not that smart or they don't cover their tracks that well.<sup>53</sup>

Under the Sarbanes-Oxley Act and the PCAOB's implementing regulations, *any* violation of laws, rules or policies by individual auditors or firms detected during inspections of selected audit and review engagements will be identified in a written report and may be handed over to the

47 SEC Chairman Harvey L. Pitt, Public Statement by SEC Chairman: Regulation of the Accounting Profession (Jan. 17, 2002) (transcript available at <http://www.sec.gov/news/speech/spch535.htm>).

48 SEC Deputy Chief Accountant Scott A. Taub, Remarks at the University of Southern California Leventhal School of Accounting SEC and Financial Reporting Conference (May 27, 2004) (emphasis added) (transcript available at <http://www.sec.gov/news/speech/spch052704sat.htm>).

49 SEC, Report Pursuant to Section 704 of the Sarbanes-Oxley Act of 2002 (Jan. 24, 2003), available at <http://www.sec.gov/news/studies/sox704report.pdf>.

50 *Id.* at 3.

51 *Id.* at 40.

52 GAAS expressly recognizes that a properly performed and executed audit may fail to detect fraud. SAS 99, *Consideration of Fraud in a Financial Statement Audit*, explains how fraud is less likely to be detected when it involves concealment and collusion: “[A]bsolute assurance [that financial statements are free of material misstatement caused by fraud] is not attainable and thus even a properly planned and performed audit may not detect a material misstatement resulting from fraud. A material misstatement may not be detected because of the nature of audit evidence or because the characteristics of fraud as discussed above may cause the auditor to rely unknowingly on audit evidence that appears to be valid, but is, in fact, false and fraudulent.” AU Section 316.12.

53 *The Enforcer*, CFO.com (Aug. 2, 2004) (emphasis added).

SEC or other regulatory authorities and become the subject of further investigation and disciplinary proceedings.<sup>54</sup>

In 2004, in its first limited inspection reports on each of the four major accounting firms,<sup>55</sup> the Board criticized two firms for audits that lacked adequate audit evidence, including the analysis of counsel regarding contingent liabilities under FAS 5.<sup>56</sup> The PCAOB, acting as the new standard setter for public company audits, issued standards on audit documentation and on audits of internal controls over financial reporting, both of which seem to encourage more rigorous audit inquiries into areas involving legal compliance and advice of counsel.<sup>57</sup> These developments suggest that auditors will be making requests not simply for relevant factual information to back up management's representations, but also for *privileged* information in order to perform financial statement audits and reviews.<sup>58</sup>

The combination of these factors has had a direct impact on the relationship between corporations and their auditors and, in turn, on the attorney-client relationship and related protections and has caused many in both communities to regard reliance on the protections as illusory. It was in partial response to these pressures that the previously-mentioned ABA Task Force was established, and the auditor-client relationship was a principal part of the Task Force's work. In August 2006, the Task Force adopted a "Report and Recommendation to ABA House of Delegates on Audit Issues," stating that "it is both appropriate and necessary for the ABA to adopt a new resolution that directly addresses erosion of the privilege and the work product doctrine in the context of audits of financial statements."<sup>59</sup> The Task Force Report identified several examples of erosion of the protections in the audit context, as reported by corporations and their internal and external counsel, including the following:

- auditor requests for a much broader range of documents in the possession of the audited company;
- auditor requests for documents covered by the protections, even where there may be other sources of the relevant information or other ways of satisfying audit needs;
- departures from the Treaty and an increase in non-standard requests;
- expansive treatment of documents in the files of an audited company as constituting audit documentation, and thus necessary to obtain as part of the audit; and
- efforts to review protected materials unrelated to the audit of the financial statements in order to provide the internal controls certification required under the Sarbanes Oxley Act Section 404.<sup>60</sup>

In light of the comments emanating from the SEC and the criticisms levied in the PCAOB's first limited inspection reports, auditors may conclude that it would be imprudent in this climate *not* to demand access to privileged information of the sort described above. However, where auditors' demands for otherwise-protected information lead to disclosure by the company, there is a general fear that a full waiver of the protections will result. This fear is warranted. The lack of any

54 When the PCAOB believes that an act, practice or omission by a registered firm or individual auditor may violate the Sarbanes-Oxley Act, PCAOB rules or other professional standards or any securities law or regulation pertaining to audit reports or to the duties of accountants, the PCAOB may open an investigation. See PCAOB R. 5101. Such an investigation can lead to disciplinary proceedings, exposing the offending auditor or firm to penalties ranging from compulsory training and mandated quality control procedures to heavy civil fines and temporary or permanent suspension from audit practice.

55 Each of the four 2003 Limited Inspection Reports issued by the PCAOB are available at <http://www.pcaobus.org/Inspections>.

56 PCAOB, Report on 2003 Limited Inspection of Deloitte & Touche LLP (Aug. 26, 2004) at 19-20, available at [http://www.pcaobus.org/documents/Inspections/2004/Public\\_Reports/Deloitte\\_Touche.pdf](http://www.pcaobus.org/documents/Inspections/2004/Public_Reports/Deloitte_Touche.pdf); KPMG Report, *supra* note 110, at 23 n.4.

57 Auditing Standard No. 2: An Audit of Internal Control over Financial Reporting Conducted In Conjunction with an Audit of Financial Statements (PCAOB, June 2004); Auditing Standard No. 3: Audit Documentation (PCAOB, August 2004).

58 It is difficult to gauge the precise degree to which the PCAOB believes company auditors should obtain attorney-client or work product protected documentation in the course of their audits. The Board's publicly available resources only rarely reference such situations. See, e.g., PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD, REPORT ON 2004 INSPECTION OF PRICEWATERHOUSECOOPERS LLP 11 (Nov. 17, 2005) available at [http://www.pcaobus.org/inspections/public\\_reports/2005/pricewaterhousecoopers.pdf](http://www.pcaobus.org/inspections/public_reports/2005/pricewaterhousecoopers.pdf).

59 See ABA Presidential Task Force on the Attorney-Client Privilege: Report and Recommendation to ABA House of Delegates on Audit Issues (June 14, 2006), available at [http://www.abanet.org/buslaw/attorneyclient/materials/hod/0806\\_report.pdf](http://www.abanet.org/buslaw/attorneyclient/materials/hod/0806_report.pdf).

60 See *id.* at p. 6.

clear rule in federal and state jurisdictions precluding waiver is paramount to a lack of comfort that the protections will remain intact. While most courts support the argument that disclosure of work product to auditors should not waive the protection as to adversaries, the decisions are not uniform and some courts would hold that disclosure constitutes a waiver. Companies, therefore, have no guarantee that courts will protect the work product generated from internal investigations from waiver as to adversaries if these materials are disclosed to auditors. This uncertainty undermines the purpose of the privileges.<sup>61</sup>

Current public sentiment and public policy regarding corporate wrongdoing favor the bending of a company's interest in maintaining the attorney-client privilege and the protections of the work product doctrine to the demands of independent auditors. Significant case law referred to above suggests that yielding protected materials to auditors may result in a waiver. Yet, a balance can be struck between the two competing interests, as also suggested by the case law. One method of resolving the tensions between the two interests is to acknowledge the common interest between auditor and client as well as upholding the expectation that the auditors will maintain the information in confidence.

Other ways in which the issues between clients and their auditors can be dealt with have been suggested by the ABA Task Force on the Attorney-Client Privilege, including identifying the specific ways in which auditors may examine evidential matter without intruding on the area of counsel's privileged advice.<sup>62</sup>

First, with respect to tax advice and opinions, the Task Force pointed out that, under existing AICPA Standards of Field Work, "[i]f the client's support for the tax accrual of matters affecting it, including tax contingencies, is based upon an opinion issued by an outside advisor with respect to a potentially material matter, the auditor *should* obtain access to the opinion, notwithstanding potential concerns regarding attorney-client or other forms of privilege."<sup>63</sup> Before the opinion should be sought, the company must base its support for its tax position upon the opinion of counsel. It is only when the company seeks to justify its tax position on counsel's opinion that the standard calls for the auditor to have access to the opinion. Thus, in the area of tax advice and opinions from counsel, auditor's requests should be limited to those circumstances in which the opinion or advice is asserted by the company as the basis for its tax position. In most circumstances, however, it should be possible for the company to produce materials satisfying audit requirements that disclose the factual and legal bases for the tax position taken by the company without the need for inquiry by the auditor into the advice or opinion of counsel.<sup>64</sup>

Applying the same analysis to other situations, the Task Force was of the view that, with respect to litigation reserves, if the client seeks to support its litigation reserve by reference to the opinion or assessment of counsel, auditors should have access to that opinion. Otherwise, the auditor should refrain from so seeking the privileged material. Of course, the underlying information relevant to determining the proper amount of the reserve is not the subject of the protections in the first instance and therefore should not be withheld from its auditor. Furthermore, an auditor may, if necessary, seek confirmation from the client that the client's position on the litigation reserve is not inconsistent with the advice of its counsel.

1. Similarly, environmental contingencies, including those involving the future obligation to retire assets covered by FAS No. 143 and Interpretation No. 47, involve

61 For example, compare *Medinal, Ltd. v. Boston Scientific Group*, 214 F.R.D. 113, 115 (S.D.N.Y. 2002) (the defendant engaged counsel to perform an investigation and to report the results of the investigation to a Special Litigation Committee ("SLC") of the Board and to the defendant's auditors in connection with their audit of loss contingency reserves. Held: because of the "independent" role of the auditor, the auditor and client did not share "common interests," thus, the disclosure waived the work product protection) with *Merrill Lynch & Co. v. Allegheny Energy, Inc.*, 229 F.R.D. 441 (S.D.N.Y. 2004) (plaintiff's disclosure to its auditors of internal investigative reports did not waive attorney work product protection, notwithstanding the independence of outside auditors, because "[a] business and its auditor can and should be aligned insofar as they both seek to prevent, detect, and root out corporate fraud"; the attorney-client privilege and work product confidentiality help to further legal compliance, thus preventing corporate fraud, much as independent auditing does; hence, the interests of the company's attorneys and its auditors were aligned).

62 See *supra* note 8, at 7-11.

63 Emphasis added.

64 Tax advice and tax opinions inherently involve legal analysis and determinations intended to be covered by the protections. These protections are essential to permit taxpayers to receive effective tax advice and to have the benefits of an adversarial system in controversies with the government. Moreover, tax matters are uniquely within the expertise of accounting firms, thus reducing their need to obtain protected analyses of counsel.



considerations similar to those with respect to tax matters. Questions relating to the proper accounting for environmental contingencies can involve legal determinations, such as whether environmental laws require remediation or taking an asset out of service and the expected timing of such actions. Counsel's assessment of these matters should only be sought by the auditor if the client justifies its position on the contingency by use of counsel's advice.

Finally, the PCAOB should formally acknowledge that the ABA Statement of Policy – the “Treaty” – has worked well for 30 years as a practical approach to preserving the protections in the context of communications directly between lawyers and auditors of companies, and should be reaffirmed. The PCAOB could also issue interpretive guidance, with approval by the SEC, advising that an auditor is generally expected to obtain adequate evidence to support its conclusions without demanding information protected by the attorney-client privilege or work product doctrine. An auditor should only require such information if it determines that there are no other sources from which it can fulfill its professional obligations.<sup>65</sup>

By issuing such guidance, the PCAOB and the SEC would acknowledge and support the compelling public interest served by protecting the confidentiality of attorney-client communications and work product, as did the AICPA in issuing its guidance that auditors need not seek access to a client's privileged information under SAS 12, beyond the Treaty, in order to audit litigation contingency reserves.

Finally, the PCAOB should promulgate guidance that an auditor does not violate independence standards by entering into a written agreement with a client providing for the confidential treatment of client information provided to the auditor, subject to the auditor's professional obligation to cooperate with the PCAOB and other oversight bodies.

By issuing such guidance, the PCAOB would further assist companies that must make their case in court for non-waiver by allowing auditors to enter into confidentiality agreements with clients. Confidentiality agreements have been crucial in the handful of decisions finding non-waiver despite disclosure of work product to government investigators.<sup>66</sup>

## VII. PROPOSED CHANGES TO THE FEDERAL RULES OF EVIDENCE

In early 2006, at the request of the Chairman of the House Committee on the Judiciary, the Judicial Conference Advisory Committee on Evidence Rules (the “Advisory Committee”) drafted a proposed rule of evidence, rule 502, entitled “Attorney-Client Privilege and Work Product; Waiver by Disclosure,” governing issues such as inadvertent disclosure, selective waiver, subject matter waiver, and the binding effect of confidentiality orders.

One particular section of the proposed rule – section 502(b)(3) – dealt with the so-called “selective waiver” doctrine by which a company first produces protected materials to a governmental authority, and later seeks to protect production to a third party in litigation.<sup>67</sup> Courts are routinely asked to opine on the question of whether that company can still invoke the confidentiality protections of the privilege and, usually but not always, rule that a company cannot

<sup>65</sup> This approach is consistent with the AICPA's 1977 guidance regarding SAS 12, discussed above.

<sup>66</sup> See, e.g., *Saino v. McKesson HBOC, Inc.*, 2002 WL 31657622, at \*6, 11 (Del. Ch. Ct. Nov. 13, 2002) (“[P]ublic policy seems to mandate that courts continue to protect the confidentially disclosed work product in order to encourage corporations to comply with law enforcement agencies.”); *Maruzen Co., Ltd. v. HSBC USA, Inc.*, 2002 WL 1628782, at \*2 (S.D.N.Y. June 23, 2002) (denying motion to compel because defendants had confidentiality agreements with U.S. Attorney's Office to whom documents were disclosed (citing *In re Steinhardt Partners, L.P.*, 9 F.3d 230, 236 (2nd Cir. 1993)).

<sup>67</sup> In a related development, on October 16, 2006, President Bush signed the so-called Regulatory Relief Bill, which contains a subsection designed to preserve the attorney/client privilege for disclosures “by any person of any information to any Federal banking agency, State bank supervisor, or foreign banking authority.” The specific provision dealing with nonwaiver of privileges provides that “The submission by any person of any information to any Federal banking agency, State bank supervisor, or foreign banking authority for any purpose in the course of any supervisory or regulatory process of such agency, supervisor, or authority shall not be construed as waiving, destroying, or otherwise affecting any privilege such person may claim with respect to such information under Federal or State law as to any person or entity other than such agency, supervisor, or authority.”

selectively waive as to the government and successfully maintain the privilege as to third parties.<sup>68</sup> It is as a result of such inconsistent court rulings that the “selective waiver” portion of the Rule was proposed to create a measure of certainty by providing that disclosure of protected material to a local, state, or federal investigating authority would not constitute a general waiver of either the attorney-client privilege or work-product protection,<sup>69</sup> and it would “further[] the important policy of cooperation with governmental agencies, and maximize[] the effectiveness and efficiency of governmental investigations.”<sup>70</sup>

However, concerns were raised that adopting the rule in the current “culture of waiver” environment may permanently prevent companies from ever being able, practically, assert a privilege again in governmental investigations, as well as whether a federally-enacted selective waiver rule would be effective in all instances in the courts of the various United States.<sup>71</sup> During the remainder of 2006 and early 2007, the Advisory Committee held additional hearings, including on January 29, 2007, at which all witness except one opposed the selective waiver portion of the proposed rule. In April 2007, a final Committee version of Rule 502 was adopted, in which provisions dealing with selective waiver were separately broken out and placed into a separate statute for Congress to consider, with the Committee taking no position on the merits of selective waiver. As separately redrafted, the provision states that “In a federal [or state] proceeding, the disclosure of a communication or information protected by the attorney client privilege or as work product – when made for any purpose to a federal office or agency in the course of any regulatory, investigative, or enforcement process – does not waive the privilege or work-product protection in favor of any person or entity other than a [the] federal office or agency.”<sup>72</sup>

Following review and presumably approval by the Standing Rules Committee on Rules of Practice and Procedure, the proposed Rule will be considered by the appropriate Congressional Committees, where it is possible that the selective waiver provision of Rule 502 will be given further consideration.

## VIII. CONCLUSION

As should be clear from the foregoing, the attorney-client privilege and work product doctrine, long the bedrock of the system by which we administer the system of justice in this country, have been put under severe strain, through a combination of governmental policies designed to root out corporate criminality, and the expected function of independent auditors, upon whose work, integrity, and independence, much of the financial disclosure system in the United States relies. Many initiatives have been undertaken to preserve the most important aspects of the protections, but whether any of them can be is unsettled at this time.

68 Compare (as to the same materials produced to the SEC by McKesson HBOC), *Saito v. McKesson HBOC, Inc.*, 2002 Del. Ch. LEXIS 139 [2002 WL 31657622 at p. \*11 (Del. Ch. Ct. Nov. 13, 2002) (“[P]ublic policy seems to mandate that courts continue to protect the confidentially disclosed work product in order to encourage corporations to comply with law enforcement agencies.”), with *McKesson HBOC, Inc. v. Superior Court*, 115 Cal. App. 4th 1229, 1241 (Cal. Ct. App. 2004) (“Given the Legislature’s expressed desire to control evidentiary privileges and protections, adoption of the selective waiver theory should come from that body. We agree with the trial court that under California law, McKesson waived the work product protection for the audit committee report and the interview memoranda.”); see also *In re McKesson HBOC Securities Litig.*, 2005 U.S. Dist. LEXIS 7098, upholding the selective waiver doctrine on behalf of McKesson HBOC.

69 See Committee Note, at 6-7 (*infra*, fn 20).

70 *Id.*

71 Compare Broun, “Memo on the Impact of the Draft Rule 502 on Waiver of Privilege in a State Action,” Capra and Broun, Memorandum to Advisory Committee on Evidence Rules, April 24-25, 2006, at 17-23, with Comment Letter, Corporations Committee of the Business Law Section of the State Bar of California, at 2.

72 See <http://www.josephnyc.com/blog/> and <http://wolfs2cents.wordpress.com/2007/04/14/advisory-committee-on-evidence-rules-completes-meetings-on-proposed-rule-502/>

# PROTECTION OF THE ATTORNEY-CLIENT PRIVILEGE IN CRIMINAL INVESTIGATIONS

---

*Barry M. Sabin and Matthew R. Lewis*  
*Department of Justice - Criminal Division*  
*Washington, D.C.*

## PROTECTION OF THE ATTORNEY-CLIENT PRIVILEGE IN CRIMINAL INVESTIGATIONS<sup>1</sup>

As most practitioners of criminal law are aware, the government has a number of methods it may use to collect evidence in a criminal investigation. This paper discusses these evidence-gathering tools generally and sets forth the limitations imposed or undertaken voluntarily by the government to ensure that the attorney-client privilege is protected when the various investigative tools are used. As its policies, practices, and pronouncements repeatedly demonstrate, the Department of Justice has a meaningful respect for the attorney-client privilege. Attorney General Alberto Gonzales captured and summarized this sentiment last month at the American Bar Association Annual White Collar Crime Conference when he stated, "I know from my own days in private practice the importance of the attorney-client privilege. I believe the McNulty Memo strikes the right balance in this area."<sup>2</sup> Prosecutors and agents, both at headquarter posts and in the field, share this respect as well, and also have powerful professional and practical incentives to respect and preserve the attorney-client privilege.

### I. THE GOVERNMENT HAS STRONG INCENTIVES TO PROTECT THE ATTORNEY-CLIENT PRIVILEGE

The federal government is obligated to vigorously enforce the federal criminal laws of the United States. Critics of prosecutorial practices maintain that while this is a noble and appropriate goal, prosecutorial zeal results in the practical result of minimizing the importance of the privilege and a cavalier disregard of the attorney-client privilege as an inconvenient impediment to enforcement obligations. Federal criminal prosecutors, however, have a strong incentive to protect the attorney-client privilege because it is fundamental to the American justice system, consistent with their legal obligations, their professional and ethical duties, and they are subject to rigorous institutional oversight mechanisms.

The Department of Justice's commitment to the protection of the privilege was most recently memorialized in the revised Principles of Federal Prosecution of Business Organizations (the "McNulty Memorandum") in December 2006. The McNulty Memorandum explicitly recognizes that "the attorney-client and work product protections serve an extremely important function in the U.S. legal system. The attorney-client privilege is one of the oldest and most sacrosanct privileges under U.S. law. See *Upjohn v. United States*, 449 U.S. 383, 389 (1976). As the Supreme Court has stated its purpose is to encourage full and frank communication between attorneys and their clients and thereby promote broader public interests in the observance of law and administration of justice.' *Id.* The work product doctrine also serves similarly important interests."

---

<sup>1</sup> Nothing in this document is intended to create any substantive or procedural rights, privileges, or benefits enforceable in any administrative, civil or criminal matter.

<sup>2</sup> See [www.usdoj.gov/ag/speeches/2007/ag\\_speech\\_070301.html](http://www.usdoj.gov/ag/speeches/2007/ag_speech_070301.html).

Professional obligations and Department of Justice investigatory components ensure that prosecutors, both in word and in practice, are committed to this appreciation for the attorney-client privilege as vital to our system of justice. Upon joining the Department of Justice, attorneys execute a written document that acknowledges their understanding of professional obligations and responsibilities. Each United States Attorney's Office has a Professional Responsibility Officer ("PRO") who is tasked to provide guidance, counsel and training to prosecutors on professional and ethical issues, including those related to privilege. All Department of Justice attorneys are required to have annual training on ethics issues. The Professional Responsibility Advisory Office ("PRAO") is a component at Main Justice in Washington, D.C. that provides definitive advice to Department of Justice attorneys on issues relating to professional conduct, including attorney-client privilege issues. Pursuant to Title 28, United States Code, Section 530B, commonly known as the "McDade Amendment," Department of Justice attorneys are legally bound by State laws and rules, and Federal court rules, governing attorneys. *See* 28 U.S.C. Section 530B ("An attorney for the Government shall be subject to State laws and rules, and local Federal court rules, governing attorneys in each State where such attorney engages in that attorney's duties, to the same extent and in the same manner as other attorneys in that State.").

There are also enforcement and oversight components in the Justice Department that ensure that prosecutors abide by these professional duties and responsibilities. The Office of Professional Responsibility ("OPR") is "responsible for investigating allegations of misconduct involving Department attorneys that relate to the exercise of their authority to investigate, litigate or provide legal advice, as well as allegations of misconduct by law enforcement personnel when they are related to allegations of attorney misconduct within the jurisdiction of OPR."<sup>3</sup> The Department of Justice's Office of Inspector General "conducts independent investigations, audits, inspections and special reviews of United States Department of Justice personnel and programs to detect and deter waste, fraud, abuse and misconduct and to promote integrity, economy, efficiency and effectiveness in Department of Justice operations."<sup>4</sup> There have also been Congressional oversight hearings in both the Senate and the House of Representatives during the past few months regarding the manner in which the Department of Justice has undertaken its practices regarding the attorney-client privilege. Prosecutors are mindful of these rigorous oversight mechanisms. Pursuant to these professional obligations, as a general proposition, prosecutors take affirmative steps to ensure that the privilege is protected.

As a practical matter, prosecutors also realize that it is often counter-productive to intrude upon the privilege because any intrusion – or even arguable intrusion – will likely result in lengthy pre-trial motion practice that will consume scarce resources and divert attention from the true issue in the case – the guilt or innocence of the defendant. Prosecutors know that issues involving the attorney-client privilege are sensitive and that defense counsel is likely to litigate fully any supposed intrusion on the privilege. The threat of a lengthy pretrial process, including the difficulties associated with *Kastigar*-like evidentiary hearings, focusing upon the actions of government actors rather than the indicted defendants, serves as a strong practical deterrent to government conduct. *See Kastigar v. United States*, 406 U.S. 441 (1972); *In Re Grand Jury Subpoenas*, 454 F.3d 511, 517 (6th Cir. 2006) ("Indeed, the government concedes that the leaking of privileged materials to investigators would raise the specter of *Kastigar*-like evidentiary hearings, and argues that it would therefore act conservatively, and err on the side of caution, in assessing the existence of privilege and in screening privileged documents from investigations."). Thus, in addition to professional motivations relating to the importance of the attorney-client privilege, there are practical considerations that motivate prosecutors to respect the attorney-client privilege.

<sup>3</sup> *See* [www.usdoj.gov/opr](http://www.usdoj.gov/opr) ("The objective of OPR is to ensure that Department of Justice attorneys continue to perform their duties in accordance with the high professional standards expected of the Nation's principal law enforcement agency.") OPR publishes an annual report relating to the investigations that were conducted during that year.

<sup>4</sup> *See* [www.usdoj.gov/oig](http://www.usdoj.gov/oig). The Inspector General publishes its reports, including their findings and recommendations regarding professional conduct.

## II. SOME EVIDENCE-GATHERING TECHNIQUES AVAILABLE TO THE GOVERNMENT IN A CRIMINAL INVESTIGATION

Federal criminal law enforcement authorities have many evidence-gathering tools at their disposal when conducting a criminal investigation. These tools include: (1) search warrants, (2) grand jury and administrative subpoenas, (3) wiretaps, (4) consent searches, and (5) voluntary disclosures, as well as others.<sup>5</sup> Which evidence gathering technique a prosecutor chooses to use in a particular criminal investigation is generally governed by the facts and circumstances of the the particular matter, legal limitations, efficiency, and resources, as well as the prosecutor's experience or familiarity with a particular technique.

The use of these various tools, as well as other protective actions such as Special Administrative Measures (SAMs), which are also discussed in this paper, potentially impacts the attorney-client privilege. However, limitations are in place to protect the attorney-client privilege in connection with each of these investigative tools, as well as the use of SAMs. These limitations are judicially imposed in many cases. In some cases, statutory language provides some basis to prevent the collection of privileged information. In others, the Department of Justice has voluntarily enacted guidelines to protect the privilege. These limitations are discussed more fully below.

### A. Protection of the Attorney-Client Privilege in Connection with Subpoenas

Subpoenas may seek the production of documents, testimony, or both. If a subpoena seeks testimony, the recipient may generally refuse to answer questions that would require the disclosure of attorney-client communications. This privilege is subject to the crime-fraud exception, which applies if the attorney-client communication is intended to further a crime or a fraud. Under this exception, "the attorney-client privilege is waived when the client used the attorney-client relationship to engage in ongoing fraud rather than to defend against past misconduct." *In re Special September 1978 Grand Jury (II)*, 640 F.2d 49, 59 (7th Cir. 1980). *See also United States v. Zolin*, 491 U.S. 554, 568-570 (1989)(permitting *in camera* review of documents alleged to come within crime-fraud exception to attorney-client privilege); *In Re Grand Jury*, 475 F.3d 1299 (D.C.Cir. 2007)(government established prima facie case that crime-fraud exception applied to any documents claimed to be within attorney-client privilege).

In the case of subpoenas seeking the production of documents, it is well-established that the recipient of the subpoena is entitled to withhold materials covered by the attorney-client privilege and is not required to produce those materials to the grand jury.<sup>6</sup> *See* Fed. R. Evid. 1101(c) ("The rule with respect to privileges applies at all stages of all actions, cases, and proceedings."); *Swidler & Berlin v. United States*, 524 U.S. 399, 410-11 (1998) (holding that attorney-client privilege may be asserted in response to grand jury subpoena even after client has died).<sup>7</sup> Of course, even when the privilege is available, the party asserting the privilege has the burden of making a prima facie showing that the privilege applies and must produce a privilege log that establishes a basis for the assertion of the privilege. *See In re Grand Jury Investigation*, 974 F.2d 1068, 1070-71 (9th Cir. 1992). *See also In re Grand Jury Subpoenas*, 454 F.3d 511, 515 (6th Cir. 2006)(discussing use of logs as a "fairly standard practice" by which law firms conduct privilege reviews).

In addition to these judicially imposed limitations, Department of Justice guidelines prohibit the issuance of a subpoena to an attorney seeking information relating to the representation of a client or the fees paid by such client without the prior approval of the Assistant Attorney General

5 This paper does not address certain issues relating to technological, international, or classified evidence gathering that might also implicate the attorney-client privilege but which are beyond the scope of this paper. This paper also does not address issues emanating from the work product doctrine.

6 Not every fact relating to an attorney-client relationship is privileged. For example, the fact that a lawyer has communicated with his or her attorney is generally not privileged. *In re Walsh*, 623 F.2d 489, 495 (7th Cir. 1980). The specific identity of a client, or his physical characteristics, are generally not privileged either. *Id.*; *In re Grand Jury Proceedings (United States v. Jones)*, 517 F.2d 666, 670 (5th Cir. 1975).

7 It is also beyond the scope of this document to address whether a government lawyer may assert the privilege to prevent the disclosure of evidence to a grand jury. *Compare In re A Witness Before the Special Grand Jury 2000-2*, 288 F.3d 289, 290 (7th Cir. 2002)(holding that a state lawyer may not "refuse, on the basis of the attorney-client privilege, to disclose communications with a state officeholder when faced with a grand jury subpoena") *with In re Grand Jury Investigation*, 399 F.3d 527, 535 (2d Cir. 2005)(holding that a state governor's office could invoke the attorney-client privilege in response to federal grand jury inquiries into the private conversations between the office's former chief legal counsel and the governor and staff members).

for the Criminal Division, in Washington, D.C. United States Attorneys' Manual ("USAM") 9-143.410. The request must provide, among other things, that: the information sought is not protected by a valid claim of privilege, that all reasonable attempts to obtain the information from alternative sources have proven unsuccessful, that the material is reasonably needed and is not peripheral or speculative, and the subpoena must be narrowly drawn. *Id.*

### **B. The Protection of the Attorney-Client Privilege in Connection with Electronic Surveillance**

In the case of electronic surveillance pursuant to a court-approved order under Title III,<sup>8</sup> the surveillance must "be conducted in such a way as to minimize the interception of communications not otherwise subject to interception." 18 U.S.C. 2518(5). Courts have interpreted this limitation to require law enforcement agents to minimize the interception of attorney-client privileged communications, even if those communications might otherwise be within the scope of the surveillance order. *See, e.g., United States v. Harrelson*, 754 F.2d 1153, 1169 (5th Cir. 1985) ("Section 2518(5) requires the government to minimize the interception of privileged communications."); *United States v. Chagra*, 754 F.2d 1181, 1182 (5th Cir. 1985) (same); *United States v. DePalma*, 461 F.Supp. 800, 821 (S.D.N.Y. 1978) (stating that under Title III, "once the parties have been identified and the conversation between them is determined to be nonpertinent or privileged, monitoring of the conversation must cease immediately").

### **C. Protection of the Attorney-Client Privilege in the Execution of Search Warrants**

It has been argued that an evidence-gathering technique that does not have built-in protections for the attorney-client privilege is the execution and review of items seized pursuant to a criminal search warrant. In the execution of a search warrant, agents generally arrive unannounced, secure the premises, and seize potential evidence particularly described in the warrant. Prosecutors frequently resort to the use of search warrants when there is a concern that evidence may be destroyed rather than produced in response to a less-invasive evidence gathering technique. Of course, the use of a search warrant is subject to judicial oversight. The warrant must be presented to a neutral magistrate who determines that there is probable cause for the search. Fed. R. Crim. P. 41(d).

When a search warrant is used in a criminal case, there is arguably potential for agents to seize and review materials that may contain attorney-client privileged information. This is due to the fact that documents are seized without an initial review by the owner or custodian of the documents. In addition, at least some preliminary review by the agents executing the warrant is usually necessary to determine whether the materials are within the scope of the warrant.

When privileged materials are seized by the government (and subject to the discussion below regarding imposition of "taint procedures,"), the courts have generally ordered that the materials be returned forthwith. *Klitzman, Klitzman and Gallagher v. Krut*, 744 F.2d 955, 962 (3d Cir. 1984) (affirming preliminary injunction requiring government to return documents seized from law office). Courts have not held that a search that involves the seizure of privileged information is unconstitutional. *See* Eric D. McArthur, "The Search and Seizure of Privileged Attorney-Client Communications," 72 U.Chi. L. Rev. 729, 740 (2005).

Nevertheless, the Department of Justice has voluntarily enacted guidelines and taken measures to ensure the protection of the privilege in the execution of search warrants on a practicing attorney – the situation where potential intrusion upon the privilege is greatest.<sup>9</sup> These guidelines recognize that the execution of a search warrant on an attorney may have a negative impact on the attorney-client relationship. Therefore, the following standards apply. First, prosecutors are expected to use alternative sources to gather the potential evidence, if reasonably available. Second, the United States Attorney in the district or the Assistant Attorney General must authorize the search

<sup>8</sup> 18 U.S.C. §§ 2510-2522.

<sup>9</sup> The search of a client that may result in the seizure of privileged materials would not be subject to these guidelines. However, the client would still be protected by the general requirement that privileged materials be immediately returned and any other prophylactic measures undertaken by agents requesting the search.

and must consult with the Criminal Division in Washington, D.C. Third, procedures should be implemented to ensure that privileged materials are not improperly viewed, seized, or retained during the course of a search. These procedures must include “adequate precautions to ensure that the materials are reviewed for privilege claims and that any privileged documents are returned to the attorney from whom they were seized.” Fourth, the warrant should be drawn as specifically as possible. USAM 9-13.420.

In order to comply with these guidelines and to minimize any intrusion upon the attorney-client privilege, criminal prosecutors employ several techniques. The most common is the use of a “taint team” or “privilege team” to participate in the search and review of any potentially privileged materials. The taint or privilege team is comprised of investigators who are not permanently assigned to the case. These agents and prosecutors review any potentially privileged material and separate and return any material that is privileged to the holder. Non-privileged material is passed along to the agents and prosecutors working the case. By using a taint or privilege team, the government ensures that the agents and prosecutors permanently assigned to the case are not exposed to privileged materials that could be used to develop a trial or investigation strategy or to question witnesses even if the privileged materials were actually returned. *See, e.g., United States v. Hunter*, 13 F.Supp2d 574 (D.Vt. 1998)(court approved use of government taint team). *But see In Re Search of the Scranton Housing Authority, No.04 Misc. Nos. 318-322*, 2006 WL 1722565 at \*5 (M.D.Pa. June 22, 2206)(citing various cases questioning or rejecting government taint team procedures).

In addition to privilege or taint teams of government agents, criminal prosecutors sometimes apply for judicial review of materials, request appointment of a special master, or work with defense counsel to establish other reasonable review procedures that do not involve the court. *See United States v. Abbell*, 914 F.Supp. 519 (S.D.Fl 1995)(after seizing law firm documents through a search warrant, the government employed a taint team to determine privilege; court appointed special master to review the documents). In those instances, the potentially privileged materials collected and segregated by the taint or privilege team may never be reviewed by the governmental personnel conducting the investigation, as well as the public or non-government personnel.

Moreover, even though these Department of Justice guidelines only apply to searches of attorney-offices, similar procedures are often invoked whenever agents and/or prosecutors believe there is a significant likelihood that privileged materials will fall within the scope of a search warrant. *See generally United States v. SDI Future Health, Inc.*, 464 F.Supp.2d 1027 (D. Nev. 2006)(discussing strengths and weaknesses of government taint teams in criminal investigations).

#### **D. The Protection of the Privilege in Voluntary Disclosures.**

Not all investigation techniques require compulsory means such as those discussed above. Criminal investigations often involve voluntary disclosure of information through interviews with witnesses, voluntary production of documents, or consensual monitoring, among others.

In the area of voluntary disclosure, the government has also taken steps to ensure that attorney-client privileged material is protected. Voluntary disclosure for purposes of this paper includes both privileged material produced by an individual or a company without a request from the government and material that is produced in response to a government request, but without compulsory process. In both cases, the production is voluntary because the government has no method to compel disclosure if the holder of the privilege declines to produce the material.

Some may legitimately ask why a company or individual would ever choose to waive the privilege if not compelled to do so. The short answer is that many companies choose to cooperate with a government investigation by turning over privileged materials because the companies realize that this course of action is often the best way to put the entire scenario behind it and to get back to business.

In December 2006, the Department issued the McNulty Memorandum discussed briefly on a previous page. The changes included in the McNulty Memorandum are designed to ensure that voluntary waivers of the privilege by a company are properly rewarded while also ensuring that corporations are not improperly pressured to waive the privilege. The McNulty Memorandum emphasizes that the waiver of the attorney-client privilege is not a prerequisite to a finding that a company has cooperated in the government's investigation. The government has underscored that waivers should not be routinely requested, that if the company desires to cooperate it should inform the government of the facts and the wrongdoers and if the company can undertake cooperation without waiving privilege, then the Justice Department will not seek to request a corporation consider a voluntary waiver. The McNulty Memorandum implements an incremental, narrowly tailored approach to seek to obtain privileged materials only when there is a legitimate need to obtain those materials.

Before formulating a request for privileged materials, a prosecutor must establish a legitimate need for the privileged material. A legitimate need is established by a four-factor test: (1) the likelihood and degree to which the privileged information will benefit the government's investigation; (2) whether the information sought can be obtained in a timely and complete fashion by using alternative means that do not require waiver; (3) the completeness of the voluntary disclosure already provided; and (4) collateral consequences to a corporation of a waiver. This test is significant and the McNulty Memorandum provides that this test is not satisfied if a waiver is "merely desirable or convenient."

If a prosecutor satisfies the legitimate needs test, he or she may request potentially privileged information. The prosecutor must first seek "purely factual information" (which may or may not be privileged). Purely factual information includes information such as copies of key documents, chronologies or organization charts created by counsel, or witness statements containing purely factual information. Any such request must be narrowly tailored, in writing, and be approved by the United States Attorney in consultation with the Assistant Attorney General for the Criminal Division. If purely factual information is insufficient and a legitimate need for attorney advice or opinion is necessary, a prosecutor may request such information only after receiving written authorization from the Deputy Attorney General, the second highest ranking official at the Department of Justice. The only exception to this rule is when the company is asserting an advice of counsel defense or if the crime-fraud exception applies.<sup>10</sup>

However, even if a prosecutor obtains the necessary authorization to request a waiver, the company is never compelled to produce the requested waiver and materials. Moreover, a prosecutor cannot even consider a company's refusal to provide advice or opinion material when deciding whether or not the company should be charged. A prosecutor may consider the company's refusal to provide purely factual information, but this is only one sub-factor in a nine factor test used to determine whether a company should be charged.<sup>11</sup>

### **E. The Protection of the Privilege in Connection with Special Administrative Measures**

The Code of Federal Regulations allows the Bureau of Prisons to issue Special Administrative Measures (SAMs) through which the federal prison may limit some activities such as phone calls, correspondence, visitation, and participation in media interviews. SAMs are potentially permissible when "reasonably necessary to prevent disclosure of classified information" or to "protect persons against the risk of acts of violence or terrorism." 28 C.F.R. §§ 501.2(a), 501.3(a). Usually, even if there is a threat of bodily harm, communications between an inmate and his or her attorneys remain unfettered. However, in very limited circumstances, when the SAMs are intended to protect persons against the risk of acts of violence or terrorism, the Attorney General may also authorize a

<sup>10</sup> In those cases, the prosecutor must obtain the level of approval required for purely factual information.

<sup>11</sup> The factors articulated in the McNulty Memo that prosecutors must consider in assessing corporate liability include: seriousness of the offense, pervasiveness of wrongdoing in the corporation, including complicity of management, a corporation's history, timely disclosure of wrongdoing, adequacy of an existing compliance program, a corporation's remedial actions, collateral consequences to shareholders and employees, adequacy of prosecuting employees, and adequacy of remedies such as civil or regulatory actions.



warden to “provide appropriate procedures for the monitoring or review of communications between that inmate and attorneys or attorneys’ agents who are traditionally covered by the attorney-client privilege” if reasonable suspicion exists to believe that the inmate may use communications with attorneys to further or facilitate acts of terrorism. *Id.* § 501.3(d).

In these very limited circumstances, the Department follows strict procedures to ensure that the attorney-client privilege is not compromised. First, the Bureau of Prisons must “provide written notice to the inmate and to the attorneys involved, prior to the initiation of any monitoring or review” that all communications between the inmate and attorneys may be monitored. *Id.* § 501.3(d)(2). Second, the Bureau of Prisons must “employ appropriate measures to ensure that all attorney-client communications are reviewed for privileged claims and that any properly privileged materials are not retained during the course of monitoring.” *Id.* § 501.3(d)(3). The regulations then state that “[t]o protect the attorney-client privilege and to ensure that the investigation is not compromised by exposure to privileged material relating to the investigation or to defense strategy, a privilege team shall be designated, consisting of individuals not involved in the underlying investigation. The monitoring shall be conducted pursuant to procedures designed to minimize the intrusion into privileged material or conversations.” *Id.* Thus, as in Title III proceedings, the government goes to great lengths to ensure that the attorney-client privilege is protected, even when national security is potentially at risk.

### III. CONCLUSION

Prosecutors and agents have a number of investigative tools at their disposal, the use of which may have an impact on the attorney-client privilege. As addressed above, limitations have been imposed by judicial decision or by statutory enactment to protect the attorney-client privilege in connection with the use of most of these tools. In addition, the Department of Justice has established guidelines to direct prosecutors and help ensure that the attorney-client privilege is protected. Prosecutors are fully committed to the ideals reflected by the guidelines and take proactive steps to protect the privilege. By doing so, these prosecutors help ensure that justice is served.

# THE STATUS AND CONTENT OF SOLICITOR-CLIENT PRIVILEGE IN CANADA: QUESTIONS STILL UNANSWERED

---

*The Honourable Gilles Létourneau\**  
*Federal Court of Appeal*  
*Ottawa, Ontario, Canada*

## I. THE INITIAL STATUS OF SOLICITOR-CLIENT PRIVILEGE

Solicitor-client privilege had its origins in our legal system as a rule of evidence. A witness could refuse to answer questions that would reveal privileged information between lawyer and client. In *Solosky v. The Queen*,<sup>1</sup> Justice Dickson (as he then was) of the Supreme Court of Canada traced the history of the privilege to the reign of Elizabeth I in England. At that time, the privilege “stemmed from respect for the ‘oath and honour’ of the lawyer, dutybound to guard closely the secrets of his client, and was restricted in operation to an exemption from testimonial compulsion.”<sup>2</sup> The rationale for the privilege was expressed in the 19th Century by Brougham L.C. in the case of *Greenough v. Gaskell*:

The foundation of this rule is not difficult to discover. It is not (as has sometimes been said) on account of any particular importance which the law attributes to the business of legal professors, or any particular disposition to afford them protection (though certainly it may not be very easy to discover why a like privilege has been refused to others, and especially to medical advisers).

But it is out of regard to the interests of justice, which cannot be upholden, and to the administration of justice, which cannot go on without the aid of men skilled in jurisprudence, in the practice of the courts, and in those matters affecting rights and obligations which form the subject of all judicial proceedings. If the privilege did not exist at all, every one would be thrown upon his own legal resources. Deprived of all professional assistance, a man would not venture to consult any skilful person, or would only dare to tell his counsellor half his case.<sup>3</sup>

As we shall see, this policy perspective has underpinned much of the development of the rules of solicitor-client privilege.

## II. THE ACTUAL STATUS OF SOLICITOR-CLIENT PRIVILEGE

It soon became apparent that the rule was too narrow to effectively protect and enforce the privilege. For example, privileged material or information could be seized under a search warrant and

---

\* Judge of the Federal Court of Appeal of Canada. The author wishes to thank his law clerk, Howard Kislowicz, for his able assistance in researching and preparing this paper.

1 [1980] 1 S.C.R. 821 [*Solosky*].

2 *Ibid.* at 834.

3 (1833) 39 E.R. 618 at 620, cited in *Solosky, ibid.*

examined by law enforcement agents. Derivative evidence from such examination could be used in aid of prosecution. The privilege therefore evolved from a rule of evidence to become a substantive right. As Justice Dickson noted in 1979:

Recent case law has taken the traditional doctrine of privilege and placed it on a new plane. Privilege is no longer regarded merely as a rule of evidence which acts as a shield to prevent privileged materials from being tendered in evidence in a court room. The courts, unwilling to so restrict the concept, have extended its application well beyond those limits.<sup>4</sup>

He went on to note: “One may depart from the current concept of privilege and approach the case on the broader basis that... the right to communicate in confidence with one’s legal adviser is a fundamental civil and legal right, founded upon the unique relationship of solicitor and client.”<sup>5</sup> Indeed, in the subsequent case of *Descôteaux v. Mierzwinski*, Chief Justice Lamer of the Supreme Court of Canada expanded upon these comments in stating:

There is no denying that a person has a right to communicate with a legal adviser in all confidence, a right that is “founded upon the unique relationship of solicitor and client” (Solosky, *supra*). It is a personal and extra-patrimonial right which follows a citizen throughout his dealings with others. Like other personal, extra-patrimonial rights, it gives rise to preventive or curative remedies provided for by law, depending on the nature of the aggression threatening it or of which it was the object. Thus a lawyer who communicates a confidential communication to others without his client’s authorization could be sued by his client for damages; or a third party who had accidentally seen the contents of a lawyer’s file could be prohibited by injunction from disclosing them.<sup>6</sup>

Still more recently, the Supreme Court of Canada ruled that the protection of solicitor client privilege is a principle of fundamental justice,<sup>7</sup> and as such warrants the protection of sections 7 and 8 of the *Canadian Charter of Rights and Freedoms*.<sup>8</sup>

In truth, the word “privilege” is a misnomer since the right against disclosure has become enforceable at an early stage. For example: on a search warrant or in the course of an inspection, if the right is invoked, the material will be secured until a judicial determination is made as to the existence or not of the right in relation to the secured material.

### III. THE CONTENT OF SOLICITOR-CLIENT PRIVILEGE

Initially, the focus was on a lawyer-client relationship. The privilege applied to protect legal advice sought and obtained from a lawyer. In *Solosky*, the Supreme Court of Canada adopted the following general statement from Wigmore:

Where legal advice of any kind is sought from a professional legal adviser in his capacity as such, the communications relating to the purpose made in confidence by the client are at his instance permanently protected from disclosures by himself or by the legal adviser, except the protection be waived.<sup>9</sup>

Over time, the privilege was extended to professionals, such as psychiatrists that the lawyer would consult.

For example, in *Perron c. R.*,<sup>10</sup> the Québec Court of Appeal was faced with a situation where defence counsel had engaged the services of a psychiatrist in order to prepare a defence. The

<sup>4</sup> *Solosky, ibid.* at 836.

<sup>5</sup> *Ibid.* at 839.

<sup>6</sup> *Descôteaux v. Mierzwinski*, [1982] 1 S.C.R. 860 at 871.

<sup>7</sup> *Lavallee, Rackel & Heintz v. Canada (Attorney General); White, Ottenheimer & Baker v. Canada (Attorney General); R. v. Fink*, [2002] 3 S.C.R. 209 Paragraph 36 [*Lavallee*].

<sup>8</sup> Part I of *The Constitution Act, 1982*, being Schedule B to the *Canada Act 1982 (U.K.)*, 1982, c. 11.

<sup>9</sup> 8 Wigmore, *Evidence* (McNaughton rev. 1961) para. 2292, cited in *Solosky, supra* note 1 at 835.

<sup>10</sup> (1990) 75 C.R. (3d) 382 [*Perron*].

Court reasoned in that case that solicitor-client privilege extended to the communications made between the accused and the psychiatrist, even though defence counsel was not present at the meeting between the accused and the psychiatrist. The Court explained that when defence counsel engaged the psychiatrist, he was acting within the scope of his duties, and that it was in his client's best interests for counsel to have access to the psychiatrist's expertise in formulating a defence. Canadian courts have considered other examples of persons outside the legal profession whose evidence might be subject to solicitor-client privilege in the proper circumstances. Non-lawyers to whom solicitor-client privilege has applied include private investigators,<sup>11</sup> accountants,<sup>12</sup> adjusters,<sup>13</sup> and physicians acting as medical experts.<sup>14</sup>

Over time, Canadian courts began to recognize a distinct type of privilege called "litigation privilege," similar to the American "attorney work-product doctrine."<sup>15</sup> Litigation privilege emerged as a second component of the solicitor-client privilege, and extended to a different class of materials that might not otherwise have been found to be subject to the legal advice privilege: materials prepared by third parties, or notes kept by the lawyer recording his or her thoughts on a given case.

#### IV. LITIGATION PRIVILEGE AND LEGAL ADVICE PRIVILEGE COMPARED

The Supreme Court of Canada has recently taken the view that litigation privilege and legal advice privilege are "distinct conceptual animals and not... two branches of the same tree."<sup>16</sup> This mixed metaphor compels a comparison between the two similar but divergent privileges.

##### *Underlying rationale*

The Supreme Court of Canada pointed out that legal advice privilege and litigation privilege are distinct, but "serve a common cause: The secure and effective administration of justice according to law."<sup>17</sup> Significantly, however, litigation privilege is based on a different underlying rationale than legal advice privilege. As the Supreme Court of Canada noted in *Blank v. Canada (Minister of Justice)*,<sup>18</sup> the object of litigation privilege "is to ensure the efficacy of the adversarial process and not to promote the solicitor-client relationship."<sup>19</sup> One Canadian appellate judge (prior to being appointed to the bench) put it another way:

Litigation privilege is based upon the need for a protected area to facilitate investigation and preparation of a case for trial by the adversarial advocate. In other words, litigation privilege aims to facilitate a process (namely, the adversary process), while solicitor-client [legal advice] privilege aims to protect a relationship (namely, the confidential relationship between a lawyer and a client).<sup>20</sup>

The ambit of each privilege reflects this difference in underlying rationale. As will be seen below, legal advice privilege has come to be seen as almost absolute,<sup>21</sup> while litigation privilege is relative.<sup>22</sup>

11 See *R. v. Samra* (1998), 41 O.R. (3d) 434 at 453.

12 *Cineplex Odeon Corp. v. Canada (Minister of National Revenue, Taxation - M.N.R.)*, [1994] O.J. No. 628 Paragraph 13 (Q.L.); *R. v. Canadian Territorial Helicopters Inc.*, 2004 MBQB 140.

13 See *Lamey (Litigation Guardian of) v. Rice* (2000), 190 D.L.R. (4th) 486. But see *Accident Assurance Co. v. Chrusz* (1999), 45 O.R. (3d) 321.

14 *Cracker v. MacDonald*, [1992] N.S.J. No. 410 (Q.L.).

15 See, for example, *Samson Indian Nation and Band v. Canada*, [1995] 2 F.C. 762 (F.C.A.); *Dupont Canada Inc. v. Emballage St-Jean Ltée*, [1999] F.C.J. No. 1429 per Hugessen J., affirmed (2000) 266 N.R. 366 (F.C.A.); *Belgravia Investments Ltd. v. Canada* (2002), 220 F.T.R. 246; *Richter Gedeon Vegyészeti Gyar Rt. v. Merck & Co.* (1996), 113 F.T.R. 1; *Canadian Jewish Congress v. Canada (Minister of Employment and Immigration)*, [1995] F.C.J. No. 362; *Jesionowski v. Gorecki* (1992), 55 F.T.R. 1; *Gower v. Talko Manitoba Inc.*, 2001 MBCA 11; *Chmara v. Nguyen* (1993), 85 Man. R. (2d) 227 (Man. C.A.); *Opron Const. Co. v. Alta.* (1989), 71 Alta. L.R. (2d) 28 (Alta. C.A.); *Global Petroleum v. CBI Industries Inc. et al.* (1998), 172 N.S.R. (2d) 326 (N.S. C.A.).

16 *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 7.

17 *Ibid.* Paragraph 31.

18 *Ibid.*

19 *Ibid.* Paragraph 27.

20 R.J. Sharpe, "Claiming Privilege in the Discovery Process", in *Law in Transition: Evidence*, [1984] *Special Lect. L.S.U.C.* 163, at pp. 164-65, cited in *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 28.

21 See *Goodis v. Ontario (Ministry of Correctional Services)*, [2006] 2 S.C.R. 32, *Pritchard v. Ontario (Human Rights Commission)*, [2004] 1 S.C.R. 809 Paragraph 17, *Canada (Attorney General) v. Canada (Information Commissioner)*, 2005 FCA 199.

22 This situation has not arisen without critique. One commentator has advanced a detailed argument to treat materials subject to litigation privilege in the same manner as materials subject to legal advice privilege: see J. Douglas Wilson, "Privilege in Experts' Working Papers", (1997) 76 Can. Bar. Rev. 346.

*Ambit*

The Supreme Court of Canada has held that

solicitor-client privilege must be as close to absolute as possible to ensure public confidence and retain relevance. As such, it will only yield in certain clearly defined circumstances, and does not involve a balancing of interests on a case-by-case basis.<sup>23</sup>

The strength of legal advice privilege in Canada is exemplified in the case of *Maranda v. Richer*.<sup>24</sup> In that case, the Supreme Court of Canada held that the amount of fees paid by an accused to his defence counsel is subject to legal advice privilege. The Crown had argued successfully before the Québec Court of Appeal that the amount of fees paid by the accused was a “pure fact,” and therefore not a “communication” that would be subject to legal advice privilege. Through a search warrant, it sought production of only these amounts, and not any advice that went along with the fees collected. The Supreme Court of Canada rejected this distinction, holding that “the amount of the fees must be regarded, in itself, as information that is, as a general rule, protected by legal advice privilege.”<sup>25</sup>

Indeed, one of the most delicate situations involving the application of solicitor-client privilege is when law enforcement officials seek a warrant to search law offices. Basing itself on the principle of minimal impairment of the right of privilege, the Supreme Court of Canada articulated the following general principles to govern law office searches:

1. No search warrant can be issued with regards to documents that are known to be protected by solicitor-client privilege.
2. Before searching a law office, the investigative authorities must satisfy the issuing justice that there exists no other reasonable alternative to the search.
3. When allowing a law office to be searched, the issuing justice must be rigorously demanding so [as] to afford maximum protection of solicitor-client confidentiality.
4. Except when the warrant specifically authorizes the immediate examination, copying and seizure of an identified document, all documents in possession of a lawyer must be sealed before being examined or removed from the lawyer's possession.
5. Every effort must be made to contact the lawyer and the client at the time of the execution of the search warrant. Where the lawyer or the client cannot be contacted, a representative of the Bar should be allowed to oversee the sealing and seizure of documents.
6. The investigative officer executing the warrant should report to the justice of the peace the efforts made to contact all potential privilege holders, who should then be given a reasonable opportunity to assert a claim of privilege and, if that claim is contested, to have the issue judicially decided.
7. If notification of potential privilege holders is not possible, the lawyer who had custody of the documents seized, or another lawyer appointed either by the Law Society or by the court, should examine the documents to determine whether a claim of privilege should be asserted, and should be given a reasonable opportunity to do so.

<sup>23</sup> *R. v. McClure*, [2001] 1 S.C.R. 445 Paragraph 35.

<sup>24</sup> [2003] 3 S.C.R. 193 [*Maranda*].

<sup>25</sup> *Ibid* Paragraph 33.

8. The Attorney General may make submissions on the issue of privilege, but should not be permitted to inspect the documents beforehand. The prosecuting authority can only inspect the documents if and when it is determined by a judge that the documents are not privileged.
9. Where sealed documents are found not to be privileged, they may be used in the normal course of the investigation.
10. Where documents are found to be privileged, they are to be returned immediately to the holder of the privilege, or to a person designated by the court.<sup>26</sup>

Similar concerns have arisen more recently with respect to the execution of *Anton Piller* orders. In *Celanese Canada Inc. v. Murray Demolition Corp.*,<sup>27</sup> the Supreme Court of Canada was faced with a situation that highlights the practical difficulties encountered with the seizure of electronic information - in that case, 1400 electronic documents - that may contain information subject to solicitor-client privilege. The case shows the need for the importation of strict conditions in the court order so as to prevent the violation of solicitor client privilege at the stage of the execution of the order. Special instructions are required in the order, such as:

- The creation of identified blocks segregated into an electronic folder;
- Supervision of the process by an independent solicitor to ensure the integrity of the search;
- A term setting out the procedure for dealing with solicitor-client privilege;
- Execution of the search during normal business hours in the presence of the defendant or a responsible employee;
- Maintenance of a detailed list of all of the evidence seized;
- Filing with the court of the independent solicitor's report on execution the search.

It should be noted that prejudice is presumed from an opponent's access to relevant solicitor-client confidences. The presumption is rebuttable, however, with the burden lying on the party who obtained wrongful access to the privileged information. The viewing of material subject to solicitor-client privilege by opposing counsel not only interferes with the rights of an opposing party, it can also have dire consequences for the party who had access to the information. In *Celanese*, the Court ordered the withdrawal of the law firm whose lawyer viewed the privileged material. This not only affected the law firm's business, but also the client's right to counsel of its choice. The Court held, however, that in the circumstances this right must yield to the countervailing right of solicitor-client privilege.

The limits of solicitor-client privilege have also been challenged in the context of the *Access to Information Act*. In a recent case, the Federal Court of Canada raised the following questions:

1. Does solicitor-client privilege attach to an entire document or only to those portions of the document which provide legal advice?
2. If a document covered by solicitor-client privilege contains within it a listing of other documents, which may or may not be covered by privilege, should these be severed from the privileged document?

---

<sup>26</sup> *Lavallee*, *supra* note 7 Paragraph 49.  
<sup>27</sup> [2006] 2 S.C.R. 189.

3. What is the effect of the release of information to satisfy the Crown's constitutional obligations for disclosure in criminal prosecutions as opposed to voluntary waiver?
4. Does partial disclosure under the Access to Information Act of a document for which solicitor-client privilege is claimed amount to waiver of privilege for the entire document?<sup>28</sup>

These questions raise complex issues, not least challenging of which is the reconciliation of the near-absolute nature of solicitor-client privilege with the government's obligation to sever and disclose information which is not covered by privilege. This tension may be observed in the legislation itself. Under section 23 of the *Access to Information Act*, "The head of a government institution may refuse to disclose any record requested under this Act that contains information that is subject to solicitor-client privilege."<sup>29</sup> However, when the government actor asserts privilege over a particular record, section 25 of the *Act*, compels the government actor to "disclose any part of the record that does not contain, and can reasonably be severed from any part that contains, any such information or material."<sup>30</sup> Moreover, section 25 is specifically designated as a paramount section of the *Act*. In deciding on the relationship between these two sections, the Federal Court of Canada articulated the principle:

severance within a document under section 25 is only to be affected where it is reasonable to do so. Reasonableness requires that the severed information be capable of standing independently and that severance must not result in the release of meaningless words and phrases out of context or provide clues to the content of the exempted portions. Severance must be done bearing in mind the importance of impairing solicitor-client privilege as little as possible.<sup>31</sup>

On appeal, the Federal Court of Appeal held that the Federal Court judge misapprehended the extent of the government's duty to sever and disclose information, and that

[s]ection 25 of the Act does not require the severance from a record of material which forms part of a solicitor-client communication... It is not Parliament's intention to require the severance of material that forms part of the privileged communication by, for example, requiring the disclosure of material that would reveal the precise subject of the communication or the factual assumptions of the legal advice given or sought.<sup>32</sup>

The person seeking access sought to rely on the following statement by the Supreme Court of Canada in support of his request for disclosure of information internal to privileged communications:

The language of s. 23 is, moreover, permissive. It provides that the Minister *may* invoke the privilege. This permissive language promotes disclosure by encouraging the Minister to refrain from invoking the privilege unless it is thought necessary to do so in the public interest. And it thus supports an interpretation that favours *more* government disclosure, not *less*.<sup>33</sup>

Evans J.A., writing for the Court, noted that these words did not support the disclosure sought for two reasons: first, the Supreme Court decision dealt with litigation privilege, not legal advice privilege; second, the permissive language in section 23 "reflects the fact that solicitor-client privilege may be waived by or on behalf of the client" and it could be assumed from the government's invocation of solicitor-client privilege that it had decided that waiver was not in the public's interest.<sup>34</sup>

<sup>28</sup> *Blank v. Canada (Minister of Justice)*, 2005 FC 1551 Paragraph 21.

<sup>29</sup> *Access to Information Act*, R.S.C. 1985 c. A-1, s. 23.

<sup>30</sup> *Access to Information Act*, R.S.C. 1985 c. A-1, s. 25.

<sup>31</sup> *Blank v. Canada (Minister of Justice)*, 2005 FC 1551 Paragraph 36.

<sup>32</sup> *Canada (Minister of Justice) v. Blank*, 2007 FCA 87 Paragraphs 7, 13.

<sup>33</sup> *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 52.

<sup>34</sup> *Canada (Minister of Justice) v. Blank*, 2007 FCA 87 Paragraph 7, 13.

Another issue that arises in the access to information context is the confidentiality of records created by the government in its processing of an access request that involves privileged material. When an access request is made, the government agency that is subject to the request will generally establish a written record of communications with other government agencies in order to come to a decision regarding the disclosure of the requested materials. These records will often make direct reference to the privileged material. They may also form the basis for the government's claim that a document is privileged. It has been argued (in a case pending before the Federal Court of Appeal at the time of writing) that, by analogy, the practice referred to as "pleading in reliance" on legal advice should apply. In such cases reliance on solicitor-client privileged advice in support of a course of action may amount to an implied waiver of that privilege. The Federal Court of Appeal held that "relying on the fact of taking legal advice necessarily puts in issue the communications between the Minister's officials and her legal advisors in such a way that it would be unfair to shield those communications from disclosure."<sup>35</sup> Accordingly, so the argument goes, records established by the government in processing an access to information request should be disclosed. The analogy, however, is strained. The underlying issue in the access context is the disclosure of documents. The disclosure of communications or advice related to those documents would have the effect of disclosing the contents of the very documents at issue in the case. It is therefore appropriate for these communications and advice to remain confidential pending the determination of the matter. Further, holding otherwise might have the undesirable consequence of encouraging the government to proceed without creating written records, thereby making the solicitor-client privilege difficult, if not impossible, to defend.

The ambit of litigation privilege has not been tested as often in the Canadian appellate courts. However, it is clear that the concerns raised by litigation privilege are not as fundamental as those raised by legal advice privilege. To reflect the differing policy considerations, litigation privilege is limited in an important way: not every document created in the events leading up to a litigated matter will be covered by litigation privilege. As Létourneau J.A. of the Federal Court of Appeal noted, litigation privilege "came to protect materials brought into existence for the dominant purpose of pending or anticipated litigation."<sup>36</sup> (emphasis added) In other words, if the dominant purpose for the creation of the material was something other than pending or anticipated litigation, the privilege will not apply.

### *Exceptions*

In the context of this robust approach to solicitor-client privilege, the Supreme Court of Canada has recognized limits and exceptions to solicitor-client privilege. Not all communications made to a lawyer will be covered by legal advice privilege. In order to attract the protection of the privilege, a communication must be made to a lawyer or his/her associates in their professional capacity, *i.e.* with the intention of seeking legal advice.<sup>37</sup> Moreover, the client must intend for the communication to be confidential when the communication is made.<sup>38</sup> Additionally, acts carried out by the client can constitute waiver of the privilege. Thus, the Québec Court of Appeal held that a legal opinion drafted by a lawyer lost its privileged status when it was passed on not only to the client, but, with the client's consent, also to the client's trustee in bankruptcy.<sup>39</sup>

Even when communications are made to a lawyer for the purpose of obtaining legal advice and the communications are intended to be confidential, certain exceptions to legal advice privilege have been recognized. These exceptions include what may be termed the "unlawful purpose" exception, the "innocence at stake" exception, and the "public safety" exception.

The unlawful purpose exception came to bear in *Descôteaux v. Mierzwinski*. The Court was faced with a situation wherein an applicant for legal aid was suspected of misrepresenting his income on his application form in order to qualify for the receipt of this

<sup>35</sup> *Canada (Minister of Health) v. Apotex Inc.*, 2004 FCA 280 Paragraph 2.

<sup>36</sup> *Blank v. Canada (Minister of Justice)*, 2004 FCA 287, dissenting on a different point, *aff'd* on this point *Blank v. Canada (Minister of Justice)*, 2006 SCC 39.

<sup>37</sup> *Descôteaux v. Mierzwinski*, *supra* note 6 at 872-873; *Solosky*, *supra* note 1 at 835.

<sup>38</sup> *Solosky*, *supra* note 1 at 835.

<sup>39</sup> *Pfeiffer v. Javicoli*, [1994] R.J.Q. 1.



government service. If true, this allegation would have constituted an indictable offence. Law enforcement officials sought access to the forms filled out by the applicant in order to prosecute the alleged offence. The Supreme Court of Canada held that the communications made by the applicant in filling out the form were made with a view to obtaining legal advice and with the expectation of confidentiality. As the false under-reporting of income was criminal in itself, however, it fell within the exception of communications that are “in themselves criminal or that are made with a view to obtaining legal advice to facilitate the commission of a crime.”<sup>40</sup> As such, the portions of the application form in which the applicant reported his income were excepted from solicitor-client privilege, while the other information contained in the form remained privileged.

In *R. v. McClure*,<sup>41</sup> the Supreme Court dealt with the “innocence at stake” exception. In that case, Mr. McClure stood accused of certain sexual offences allegedly committed against students of a school in which he worked as a librarian. After reading of McClure’s arrest in the newspaper, a person identified in the case as J.C. made statements to the police in which he alleged that he was a victim of sexual touching at the hands of McClure. These allegations were added to the indictment against McClure. J.C. also initiated a civil suit against McClure. In the context of his criminal defence, McClure sought production of the civil litigation file kept by J.C.’s counsel, arguing that the communications made by J.C. to his lawyer could go towards establishing McClure’s innocence. In refusing to compel the civil litigation file’s production, the Supreme Court of Canada formulated a two-stage test for determining if the innocence at stake exception applies. Before the two-stage analysis is considered, the accused “must establish that the information he is seeking in the solicitor-client file is not available from any other source and he is otherwise unable to raise a reasonable doubt as to his guilt in any other way.”<sup>42</sup> Once the accused overcomes this initial hurdle, the two-stage test proceeds. At the first stage, the party seeking production bears an initial burden of providing an evidentiary basis for its request. The trial judge must ask: “Is there some evidentiary basis for the claim that a solicitor-client communication exists that could raise a reasonable doubt about the guilt of the accused?”<sup>43</sup> If the trial judge is satisfied that such an evidentiary basis exists, he or she should examine the record without yet producing it to the accused. Upon examining the record, the trial judge should produce it to the accused if satisfied that something in the communication would be likely to raise a reasonable doubt about the accused’s guilt.<sup>44</sup> The difficulties in meeting these requirements again illustrate the strength of solicitor-client privilege in Canadian law. In the *McClure* case, the accused’s request was denied because he failed to provide any evidentiary basis for the claim that the information in J.C.’s litigation record could raise a reasonable doubt as to his guilt.

In *Smith v. Jones*,<sup>45</sup> the Supreme Court of Canada considered the public safety exception to solicitor-client privilege. In that case, an accused had been charged with aggravated sexual assault on a prostitute. Defence counsel referred the accused to a psychiatrist in preparing for his defence. Counsel informed the accused that the communications made to the psychiatrist would be privileged. At the interview with the psychiatrist, the accused described a detailed plan to kidnap, rape and kill prostitutes. The psychiatrist informed defence counsel that, in his opinion, the accused was dangerous and would likely commit future offences if he did not receive treatment. The accused later pled guilty to the included offence of aggravated assault. When the psychiatrist contacted defence counsel to inquire as to the status of proceedings, he was informed that his concerns about the dangerousness of the accused would not be addressed at the accused’s sentencing hearing. The psychiatrist applied to the court for a declaration that he was entitled to disclose the information in his possession in the interests of public safety. The Supreme Court held that the psychiatrist should be permitted to disclose the statements made by the accused and the psychiatrist’s opinion about them to the police and the Crown. The Supreme Court further directed that the elements of the psychiatrist’s affidavit which fell within the public safety exception to solicitor-client privilege should not be subject to a publication ban. In making its order, the Supreme Court articulated a three-step approach to determining when the public safety exception applies: “First, is there a clear risk to an identifiable person or group of persons? Second, is there a risk of serious bodily harm or death? Third, is the danger imminent?”<sup>46</sup>

40 *Descôteaux v. Mierzewski*, *supra* note 6 at 893

41 *Supra* note 23.

42 *R. v. McClure*, *supra* note 23 at 48.

43 *Ibid.*, Paragraph 52.

44 *Ibid.*, Paragraph 58.

45 [1999] 1 S.C.R. 455.

46 *Ibid.*, Paragraph 77.

These exceptions to the rule of legal advice privilege apply equally in the case of litigation privilege. As litigation privilege has been consistently seen as relating less to a fundamental right than solicitor-client privilege, there is no reason to expect that litigation privilege should not give way in the same circumstances.<sup>47</sup>

### *Time limits*

Unlike legal advice privilege, which attaches to communications unless and until the privilege is waived by the client, litigation privilege “expires with the litigation of which it was born.”<sup>48</sup> The Supreme Court of Canada explained the reason for this temporary existence as follows:

The purpose of the litigation privilege... is to create a “zone of privacy” in relation to pending or apprehended litigation. Once the litigation has ended, the privilege to which it gave rise has lost its specific and concrete purpose - and therefore its justification.<sup>49</sup>

Writing in dissent at the Federal Court of Appeal in the *Blank* case, Létourneau J.A. raised concerns with this temporal limitation as it concerned the government as a litigant:

The government is entitled to develop, in the public interest which, contrary to private companies, it must defend, a legal policy and strategy towards the conduct of these litigations. Automatic and uncontrolled access to the government lawyer’s brief, once the first litigation is over, may impede the possibility of effectively adopting and implementing such policy or strategy. It would give opponents or adversaries access to the government agency’s mind and legal strategy, the very thing that the litigation privilege is directed at preventing. It would defeat, in subsequent litigation involving the same or substantially the same cause of action, the objective pursued by the litigation privilege.<sup>50</sup>

It must be recalled that these comments arose in the context of a request made by Mr. Blank pursuant to the *Access to Information Act*.<sup>51</sup> Létourneau J.A.’s concern was motivated by the potential impediments a limited concept of litigation privilege would place on the government developing an effective and comprehensive litigation strategy, which would apply across governmental agencies.

The majority of the Federal Court of Appeal and the Supreme Court of Canada took the view that these concerns could be answered by adopting a broad definition of the term “litigation”. This broad definition would include “separate proceedings that involve the same or related parties and arise from the same or a related cause of action... Proceedings that raise issues common to the initial action and share its essential purpose would... qualify as well.”<sup>52</sup> The effect of this broad definition is to force a return to first principles when deciding whether a particular document is covered by litigation privilege. One must refer to the litigant and lawyer’s need for a zone of privacy in order for the adversarial system to function effectively. The more closely related the issues in two sets of proceedings, the stronger the case that litigation privilege should subsist for the life of both sets of proceedings. According to the Supreme Court of Canada, if a manufacturer was dealing with separate but related claims of product liability, litigation privilege would survive as long as the group of claims remained unresolved. On the other hand, the concern that a party might discover the general litigation strategy of its former adversary is not sufficient to ground a claim of litigation privilege.<sup>53</sup>

### *Legal Advice Privilege: Identifying the Client*

Because legal advice privilege attaches to communications between lawyer and client, identifying the communications to which the privilege attaches in the case of a corporate client might

47 See *Ibid.*, Paragraph 44.

48 *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 8.

49 *Ibid.*, Paragraph 34.

50 *Blank v. Canada (Minister of Justice)*, 2004 FCA 287 Paragraph 42.

51 R.S.C. 1985, c. A-1.

52 *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 39.

53 *Ibid.*, Paragraph 40-42.

not be a simple exercise. In the American experience, the question as to which individuals can communicate on behalf of a corporate client and expect their communications to be privileged has given rise to some controversy.<sup>54</sup> Canadian courts have not seen this issue as problematic, and have generally given “broad protection for confidential communications emanating from an employee, regardless of the level of his or her position in the corporate hierarchy.”<sup>55</sup> In this regard, one Canadian court has gone so far as to consider privileged the confidential communications between a corporation’s legal counsel and the employees of a subsidiary company.<sup>56</sup>

*Common concerns: Loss of privilege*

Though legal advice and litigation privilege are different in many respects, both raise similar concerns in the way in which the privilege which attaches to communications can be lost. Authors John Sopinka, Sidney H. Lederman and Alan W. Bryant identify five potential ways in which privilege can be lost: through voluntary waiver, waiver by implication, legislation, inadvertent disclosure or intercepted communications, and by reference to public policy. The authors note that, given the constitutional status now enjoyed by solicitor-client privilege, legislation purporting to curtail the privilege will be unlikely to emerge, and will be constitutionally suspect if it does.

With respect to voluntary waiver, it has been held in Canada that waiver occurs where the holder of the privilege voluntarily discloses or consents to the disclosure of any material part of a communication.<sup>57</sup> Voluntary waiver also applies to documents which have been disclosed in another jurisdiction.<sup>58</sup> In consenting to waiver, a party can selectively waive the privilege in some of the documents in a particular file while maintaining the privilege in the other documents so long as the waiver is not misleading.<sup>59</sup> Generally, disclosure to outsiders constitutes waiver of the privilege.<sup>60</sup> However, an exception is made when the “outsider” shares a common interest in the existing or anticipated litigation with the holder of the privilege.<sup>61</sup> Though Canadian courts have not ruled on the issue, English courts have held that where privilege is waived in respect of a communication, it is waived in respect of all documents relating to the acts contained in the communication.<sup>62</sup> It has also been held in England that where a party makes privileged documents available to the police to assist in a criminal investigation, such disclosure is not considered to be a waiver of the privilege in the documents with respect to the civil action for which the documents were created.<sup>63</sup>

Courts may find an implied waiver of privilege when such is required by fairness. For example, when a party relies on legal advice it received in order to justify its actions, waiver will be implied.<sup>64</sup> Similarly, the Supreme Court held that where an accused’s counsel made reference to an expert’s report in opening statements to the jury, he had waived privilege in the entire document on the accused’s behalf.<sup>65</sup> Where, however, the existence or adequacy of legal advice is not in issue, the privilege is not waived by a reference to the legal advice.<sup>66</sup>

The traditional common law rule with respect to inadvertent disclosure was established in the English case of *Calcraft v. Guest*.<sup>67</sup> The rule states that once a communication has been shared with a third party, either by inadvertence or the stealth of the third party, privilege no longer attaches and the evidence is admissible, even compellable. However, in *Ashburton (Lord) v. Pape*,<sup>68</sup> it was held that an injunction could be obtained to prevent the admission of evidence obtained through dishonesty. Canadian courts have taken the view that the judge should have discretion to decide whether privilege

54 See for example *Upjohn v. United States*, 449 U.S. 383 (1981).

55 John Sopinka, Sidney N. Lederman & Alan W. Bryant, *The Law of Evidence in Canada*, 2d ed. (Toronto: Butterworths, 1999) at 744.

56 *Mutual Life Assurance Co. of Canada v. Canada (Deputy Attorney General)* (1988), 28 C.P.C. (2d) 101 (Ont. H.C.J.).

57 *Marlborough Hotel Co. v. Parkmaster (Canada) Ltd.* (1959), 28 W.W.R. 49 (Man. C.A.).

58 *Western Assurance Co. v. Canada Life Assurance Co.* (1987), 63 O.R. (2d) 276.

59 *Transamerica Life Insurance Co. of Canada v. Canada Life Assurance Co.* (1995), 27 O.R. (3d) 291 (Gen. Div.); *Bone v. Person* (2000), 185 D.L.R. (4th) 335.

60 *Wellman v. General Crane Industries Ltd.* (1986), 20 O.A.C. 284 (C.A.).

61 *Supercam of California Ltd. v. Sovereign General Insurance Co.* (1998), 37 O.R. (3d) 597 (Gen. Div.); *General Accident Assurance Co. v. Chrusz* (1999), 45 O.R. (3d) 321 (C.A.).

62 *Donald (George) Ltd. v. Blackburn, Robson, Coates & Co. (a firm)*, [1972] 3 All E.R. 959 (Q.B.).

63 *British Coal Corp. v. Dennis Rye Ltd.* (No. 2), [1988] All E.R. 816 (C.A.).

64 *R. v. Campbell*, [1999] 1 S.C.R. 565.

65 *R. v. Stone*, [1999] 2 S.C.R. 290.

66 *Lac La Ronge Indian Band v. Canada* (1996), 6 C.P.C. (4th) 110 (Sask. Q.B.).

67 [1898] 1 Q.B. 759 (C.A.).

68 [1913] 2 Ch. 469 (C.A.).

has been lost in inadvertently disclosed documents, and should be able to take into account whether the error is excusable, whether an immediate attempt has been made to retrieve the information, and whether preservation of the privilege will cause unfairness to the opponent.<sup>69</sup> Where information is obtained by trickery, it could be argued in criminal cases that the interception of solicitor-client privileged information is an unreasonable search and seizure and thus unconstitutional. Further, the *Criminal Code*<sup>70</sup> provides at subsection 189(6) that information intercepted under judicial authorization is not admissible where it is privileged.

The public policy exceptions to privilege have been discussed above. Under these exceptions, privilege can be lost in situations where a client seeks legal advice for an unlawful purpose,<sup>71</sup> where public safety is at stake,<sup>72</sup> and where the innocence of an accused is at stake.<sup>73</sup> Furthermore, in *Re Robertson Stromberg*,<sup>74</sup> it was held that lawyers who are pursued by their law society may not benefit from solicitor-client privilege, as the public interest in the ethical practice of law outweighs any solicitor-client privilege. In such a case, the law society would not be able to disclose the information or use it for any purpose other than its investigation and proceedings.

## V. QUESTIONS STILL UNANSWERED

While the Supreme Court of Canada's decision in *Blank* provided elaboration on the concept of litigation privilege, it left significant questions unanswered. First, what is to be done with those documents gathered or assembled - though not created - for the purpose of litigation? The court specifically chose not to address this issue as it was not explicitly raised and argued. However, the Court did provide some guidance on the issue in stating that "Extending the privilege to the gathering of documents resulting from research or the exercise of skill and knowledge does appear to be more consistent with the rationale and purpose of the litigation privilege."<sup>75</sup> Indeed, it may be difficult to draw a distinction between a report drawn up by a private investigator and a document discovered by the same investigator. On the other hand, is there a distinction to be drawn between a report of an insurance claims adjuster created at the instance of a lawyer, and a series of similar reports relating to different cases copied by the lawyer in preparation for settlement negotiations? Time will tell.

Second, it is too early to tell how the broad definition of "litigation" advocated by the Supreme Court of Canada will be applied in future cases.<sup>76</sup> Given that litigation privilege applies to ongoing or reasonably apprehended litigation, does the new broader definition of litigation expand the types of cases that would be considered reasonably apprehended? This question too will likely arise in the future.

Third, new situations will inevitably arise that will test the limits of litigation privilege. In the United States, for example, several cases have come before courts where a law firm has asserted litigation privilege over reports prepared by litigation communications specialists. Lawyers consult these specialists for advice on how to handle media questions and on preparing a communications strategy. Some commentators insist that communication with the media has become an imperative for counsel, and that these specialists provide expertise necessary to the conduct of a full and effective litigation or settlement strategy.<sup>77</sup> These commentators note that American courts have divided on this issue. It remains to be seen how the question would be handled in Canada.

Fourth, questions still remain in the context of the *Access to Information Act*. Access requests could alter the usual balancing of interests in that the right of access to information is open to all Canadians,<sup>78</sup> not just adversarial parties. In the criminal law context, particularly sensitive concerns

69 See for example *Royal Bank of Canada v. Lee* (1992), 3 Alta. L.R. (3d) 187; *Airst v. Airst* (1998), 37 O.R. (3d) 654 (Gen. Div.).

70 R.S.C. 1985, c. C-46.

71 *Descôteaux v. Mierzwiński*, *supra* note 6.

72 *Smith v. Jones*, *supra* note 45.

73 *R. v. McClure*, *supra* note 23.

74 (1994), 124 Sask R. 259 (Q.B.) aff'd (1995) 122 D.L.R. (4th) 433 (C.A.)

75 *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 64.

76 See Dale E. Ives & Stephen G.A. Pitel, "Filling in the blanks for litigation privilege: *Blank v. Canada (Minister of Justice)*" (2007) 11 International Journal of Evidence & Proof 49 at 54.

77 Steven B. Hantler, Victor E. Schwartz & Phil S. Goldberg, "Extending the Privilege to Litigation Communications Specialists in the Age of Trial by Media" (2004) 13 CommLaw Conspectus 7.

78 See *Access to Information Act*, R.S.C. 1985 c. A-1, s. 4.

exist when and if the public has a right to access the Crown's brief. Létourneau J.A. listed "Public safety, the privacy interests of victims or witnesses, protection of sources and police informants, protection of Charter rights and freedoms, protection of the integrity of the administration of the criminal justice system"<sup>79</sup> as values that may need to be balanced with the public's right to disclosure of government information. He further noted two examples of cases in which deleterious effects flowed from such disclosure:

disclosure of the Crown brief in a case of alleged sexual assault resulted in the statement of a child complainant being circulated at the complainant's school, and disclosure of statements of potential Crown witnesses in penitentiary investigations ended up posted on bulletin boards to be perused by the general inmate population.<sup>80</sup>

Finally, the Supreme Court's decision in *Blank* has been criticized for stating that "litigation privilege would not in any event protect from disclosure evidence of the claimant party's abuse of process or similar blameworthy conduct."<sup>81</sup> According to Professors Ives and Pitel of the University of Western Ontario, adopting the full force of this statement would end up expanding the "unlawful purpose" exception to solicitor-client privilege in an undesirable way. These commentators advocate a more measured exception, focusing specifically on issues that go to the heart of the administration of justice, such as abuse of process or malicious prosecution, while eschewing the broader language employed by the Court referring to "similarly blameworthy conduct."<sup>82</sup> Interestingly, another commonwealth commentator has made a suggestion in the opposite direction. For Prof. Mahoney, the prime policy objectives ought to be the search for truth and the orderly resolution of disputes. He argues that these would best be served by a rule that would "[p]ermit forced disclosure of important evidence when another party demonstrates that, through no negligence on his or her part, there are not reasonable alternative means by which the evidence may be obtained."<sup>83</sup> Given the recent case law in Canada, however, this position does not seem to be in danger of becoming law in the foreseeable future.

## VI. CONCLUSION

Litigation privilege and legal advice privilege were initially covered by the same concept, which emerged as a rule of evidence. Over time, different justifications emerged for the two types of privilege, and, accordingly each has developed its own limits. It is now proper to speak of the privileges as conceptually distinct, though both are concerned with the administration of justice. It is helpful to think of legal advice privilege as a principle of fundamental justice and a fundamental right relating to the particular relationship between a lawyer and his or her client. Litigation privilege may be thought of as a necessary ballast of an adversarial system, which preserves an incentive structure favouring complete investigation on all sides of a dispute. Questions remain, however, about how expansively courts will treat the concept of litigation privilege.

<sup>79</sup> *Blank v. Canada (Minister of Justice)*, 2004 FCA 287 Paragraph 47.

<sup>80</sup> *Ibid.*

<sup>81</sup> *Blank v. Canada (Minister of Justice)*, 2006 SCC 39 Paragraph 44.

<sup>82</sup> Ives & Pitel, *supra* note 76 at 55.

<sup>83</sup> Richard Mahoney, "Reforming Litigation Privilege", (2001) 30 C.L. World Rev. 66 at 93.

# COMPARATIVE APPROACHES TO THE ATTORNEY-CLIENT PRIVILEGE IN THE US, CANADA, UK & EU

---

*Lista M. Cannon*  
*Fulbright & Jaworski International LLP*  
*London, England*

This article summarizes the general principles of legal privilege in Europe both in the context of the civil code approach to legal privilege across the European Union (“EU”) member states and as applied by the European Commission (“EC”). The main areas of discussion are as follows:

- I. Legal Privilege - The General Principles**
- II. A Sample of the EU Member States**  
*France, Germany, Italy, The Netherlands, Spain, Belgium and Sweden*
- III. The European Commission: Is there a Harmonized Rule Within the EU?**  
*The Decision of the ECJ in AM&S*  
*The Decision Pending in Akzo Nobel and Akros Chemicals*  
*The Modernisation Regulation*
- IV. Practical Implications: Managing the Risks and Cross-Border Disputes**  
*EC Investigations: What is Privileged*  
*Best Practice: Disclosure*  
*Best Practice: Privilege and Confidentiality*

## I. LEGAL PRIVILEGE - THE GENERAL PRINCIPLES

At present, Europe is striving towards making a progression to common political and economic unity throughout the primary sovereign state cultures. This is reflected in the EU legal and enforcement structure of the EC, and the in-balance between Common Law and Civil Code procedures.

Legal professional privilege at the EU level, is a judicial creation. It has been established and developed by the Court of First Instance (“CFI”) and the European Court of Justice (“ECJ”). However, the rules of privilege of the individual member states within the EU itself are governed by their respective codes of civil procedure.

It should be noted from the outset that most civil law jurisdictions do not impose a general obligation to disclose documents that adversely affect their cases, although narrow rules are available for specific applications in relation to specifically identified documents. While applications for specific disclosure are available in most jurisdictions, it is only in defined circumstances by way of describing a document precisely and justifying the application by indicating the facts a party believes each document will prove.

Generally, the approach of the civil law jurisdictions to lawyer privilege is one of professional secrecy and confidentiality. Such rights do not attach to documents but to a professional

status and obligation. Further, there are important distinctions with regard to in-house lawyers and communications with third parties which are discussed in more detail below. Unlike the common law jurisdictions, the communication itself is not privileged, but rather, the lawyer is under a duty not to disclose the information found within that communication.

The difficulties in asserting privilege often arise when a company circulates communications across jurisdictions, either voluntarily or as required by law. In relation to cross-border disputes and at the EC level, a litigant or company involved in a regulatory investigation may be obliged to disclose communications in one jurisdiction while benefiting from privilege in relation to the same communications in another jurisdiction. In these circumstances, communications or documents and advice on sensitive matters should be circulated with care, recognising the different rules and interpretations on privilege and confidentiality.

## II. A SAMPLE OF THE EU MEMBER STATES

The following is an overview of the civil code approach to privilege across some of the EU member states. The scope of privilege in each jurisdiction is identified under their respective codes together with disclosure obligations. All information in this section is sourced from the Practical Law Company, online cross border service ([www.practicallaw.com](http://www.practicallaw.com)).<sup>1</sup>

### 1. France

#### (i) *Privilege*

The relationship between a lawyer (*avocat*, admitted to the local bar) and a client is protected by professional confidentiality obligations (*articles 226-13, New Criminal Code*), which prohibit a professional who is subject to a confidentiality obligation from divulging information obtained from their client (a criminal offence unlimited in time): the lawyer/client privilege.

All correspondence (oral or written) by a lawyer in relation to a matter handled on behalf of a client, between a lawyer and a client, and between a lawyer and opposing lawyers in relation to a matter, is protected by professional confidentiality.<sup>2</sup> Correspondence protected by professional secrecy cannot be disclosed to the Court nor transmitted to clients unless marked "*official*". Importantly, lawyer to foreign lawyer correspondence across countries is not privileged, unless subject to a separate confidentiality undertaking with the foreign lawyer. Internal communications by an in-house lawyer are not privileged.

A client cannot release their lawyer from the lawyer's obligation to keep documents confidential, but are not themselves bound by the confidentiality obligation.

#### (ii) *Disclosure*

There is no process in French civil procedure that is equivalent to documentary discovery or disclosure. Parties to civil proceedings in France generally only produce the documents that they consider to support their respective cases.

### 2. Germany

#### (i) *Privilege*

The relationship between a lawyer and a client is protected by a number of professional confidentiality regulations. Without the consent of the client, a lawyer is prohibited from divulging any confidential information or documents obtained in the course of their professional activities (*section 203(1), Criminal Code*). This obligation to preserve confidentiality is mirrored by provisions

<sup>1</sup> See "*A World Tour of the Rules of Privilege*", 1 November 2006 (and updated) Resource ID: 9-205-5802.

<sup>2</sup> Exceptions: unless there is express indication to the contrary (*articles 66-5, Law of 31 December 1971*).

which give the lawyer a right to refuse to divulge such information (*sections 383 and 142(2), Civil Procedure Code, and section 53 of the Criminal Procedure Code*).

(ii) *Disclosure*

There is no duty to disclose documents to the other side, other than those upon which a party intends to rely. As discussed above, only very limited means of obtaining disclosure from the court exist.

Civil Courts do not investigate facts in litigation; the production of facts is the responsibility of the parties to establish the burden of proof. The issue of privilege and obligation of confidentiality therefore does not generally arise in the minds of lawyers in German based matters.

Generally, documents entrusted to a lawyer in their professional capacity, and which remain in their possession, are protected from disclosure (*section 97, Criminal Procedure Code*).

### 3. Italy

(i) *Privilege*

There is no recognised doctrine of privilege in Italy. However, a lawyer cannot be obliged to give evidence of any information acquired by reason of their profession, including conversations and communications with their clients (*article 200, Italian Code of Criminal Procedure*). The same applies to any document in a lawyer's possession as a result of their professional activities, if they declare in writing that the document is covered by professional confidentiality (*article 256, Italian Code of Criminal Procedure*). Lawyer-client communications held at the client's premises are also generally protected from disclosure.

(ii) *Disclosure*

There is no formal process of disclosure. The parties are merely obliged to produce their own bundle of exhibits on which they rely, to be served on both the other side and the Court.

### 4. The Netherlands

(i) *Privilege*

Those entrusted with a duty of confidence by status or by profession (such as priests, doctors, lawyers and notaries) cannot be forced to reveal confidential information (*article 843a sub 3, Dutch Act on Procedure in Civil Matters RV and article 165 sub 2b, RV*). This right to legal privilege only relates to information revealed to lawyers in their professional capacity. The Professional Conduct Rules of the Bar forbid a lawyer from testifying to facts that were revealed to them by a client in the course of the exercise of their profession, although a client can give their lawyer permission to use specific confidential information in Court.

Correspondence between Dutch lawyers is confidential in nature and cannot be used in Court, except where the client's interests require this. However, even in such a case, the prior consent of the other part or the president of the local bar is required. Lawyer-client communications held at the client's office are protected from seizure by regulatory and other investigative bodies.

(ii) *Disclosure*

Dutch law does not provide for a general duty to disclose comparable to the UK or US discovery rules. However, the Dutch law of procedure does contain a limited number of specific regulations which allow the Court to order the disclosure of specific documents.



## 5. Spain

### (i) Privilege

Lawyers (*abogados*, for whom membership of the bar is obligatory) must keep confidential all facts and matters that they come to know through the conduct of their professional obligations (*article 542, Law of Judicial Authority*). This is reinforced by the imposition on lawyers of a duty not to disclose facts and documents that have come into their possession as a result of their professional activities (*Spanish Professional Conduct Code, (June 2000) and General Statute for Spanish Lawyers*).

Clients may not release their solicitors from this duty, although they are not bound by it themselves. However, relevant documents in the client's possession continue to benefit from confidentiality and do not have to be disclosed to investigative bodies.

### (ii) Disclosure

Disclosure must be made of the documents that a party intends to use to support its own case only.

## 6. Belgium

### (i) Privilege

As is the case in The Netherlands, those entrusted with a duty of confidence by status or by profession such as lawyers, and doctors, cannot reveal confidential information except where they are called to give evidence in legal proceedings or where the law requires them to disclose the information in question (*article 458, Belgian Criminal Code, 1867*). This concept is referred to as "professional confidentiality".

Correspondence between a lawyer and client is confidential by nature. Even the client may not produce correspondence from their lawyer marked as confidential unless the lawyer consents. However, correspondence previously marked by the lawyer as non-confidential may be used in Court.

Correspondence between Belgian lawyers is also confidential in principle and cannot be used in evidence (*article 444, Judicial Code*). Some correspondence between lawyers will be classified as "official" and can be produced in Court. The relevant Professional Conduct Rules determine how the distinction should be made.

### (ii) Disclosure

There is no formal process of disclosure. The parties are merely obliged to produce their own bundle of exhibits on which they rely, to be served on both the other side and the Court.

## 7. Sweden

### (i) Privilege

The concept of legal professional privilege is recognised, although it is limited and depends on the identity of the lawyer.

Legal privilege is primarily an exception to the general obligation to give testimony provided by the Swedish Procedural Code (SPC). Swedish *advokats* (members of the Swedish Bar Association) and their assistants have a right to legal privilege, which protects all confidential information gained by them in the provision of legal services generally (*chapter 36, section 5, SPC*). However, legal privilege available to non-advocate trial lawyers is limited to protecting only confidential client communications entrusted to the lawyer for the purposes of the litigation.

Investigative authorities are not entitled to seize lawyer-client communications held at the client's premises. This right may be overridden where the examination is authorised by law or the client consents to the disclosure.

(ii) *Disclosure*

There is no concept of disclosure of documents in Swedish law, although during the preparation for trial each party must submit all documents it wishes to present as evidence. A party must also indicate what additional items of written evidence it is holding if asked by the opposing party.

### III. THE EUROPEAN COMMISSION: IS THERE A HARMONIZED RULE WITHIN THE EU?

The EC is the executive arm of the EU which was created to ensure compliance with EU law. The EC rules on legal privilege (known as attorney-client privilege in the US), are in considerable contrast with those of the EU member states, resulting in an inconsistent application of legal privilege within the EU.

The EC undertakes investigations into anti-competitive practices: cartels, monopolies and intellectual property rights. Inquiries and investigations are also often in conjunction with the US Securities and Exchange Commission, US Department of Justice and other EU Agencies.

#### 1. The Decision of the ECJ in *AM&S*

Neither *Articles 81 and 82 of the EC Treaty*,<sup>3</sup> nor any of the regulations implementing them, contain any provisions in relation to legal privilege. The principles governing legal privilege have largely been developed through the case law of the ECJ. The *AM&S*<sup>4</sup> case established the principle that former *Regulation 17 of the EC Treaty*,<sup>5</sup> which sets out the rules implementing *Articles 81 and 82*, must be interpreted as protecting the confidentiality of written communications between lawyer and client.<sup>6</sup>

The *AM&S* principle is subject to the following two conditions:

- (i) The communications must be made for the purpose and in the interests of the client's right of defence in connection with the EC's investigation.
- (ii) The communications must emanate from independent lawyers qualified to practise in a member state within the European Economic Area ("EEA"). Meaning, the privilege of in-house and non-EEA qualified lawyers is not respected.

The second condition as held in *AM&S* was broadened in the *Hilti*<sup>7</sup> case, where it was held that an in-house lawyer within the EU who reports an external lawyers advice can in fact claim privilege on that advice, provided they do not alter the advice (i.e. provide a summary of the advice). The ECJ's narrow interpretation of legal professional privilege in *AM&S* excludes the advice of in-house lawyers because they are not considered to be independent. Further, while some in-house lawyers are subject to professional conduct rules, the majority are not.

The position taken by the ECJ is in contrast with that of some of the EU member states, where correspondence with an in-house lawyer tends to be protected if that lawyer is a member of a professional association.<sup>8</sup>

<sup>3</sup> In the EU, competition law is regulated by Articles 81 and 82 of the EC Treaty.

<sup>4</sup> *AM&S v Commission* [1982] ECR 1575.

<sup>5</sup> Replaced by *Regulation 1/2003*.

<sup>6</sup> *Regulation 17* has since been replaced by *Regulation 1/2003*, the Modernisation Regulation, but the position remains the same.

<sup>7</sup> *Hilti v Commission* Case T-30/89 [1990] ECR 163.

<sup>8</sup> Note exceptions include Member States where in-house lawyers cannot be members of the bar (e.g. France, Italy and Sweden).

EU competition law applies only where there is an effect on trade between member states and exists alongside national EU member state competition laws. However, when a competition investigation involves multiple member states, the EC will prevail. The practical effect of this is that it is possible for an in-house lawyer to be afforded privilege in a UK competition authority investigation but not to be afforded privilege if the same investigation were being carried out by the EC.

## 2. The Decision Pending in *Akzo Nobel* and *Akros Chemicals*

The status of communications from in-house lawyers is currently pending before the EU's Court of First Instance ("CFI"),<sup>9</sup> a decision which many hope will further broaden the rule set out by the ECJ in *AM&S*.

The *Akzo Nobel* case followed a dawn raid by the Commission on Akzo Nobel, during which the company argued that certain communications, including some from in-house lawyer, should be regarded as privileged. Privilege was claimed over two sets of documents. The first envelope contained documents prepared for a competition law compliance programme and the second envelope contained drafts of the first set of documents together with correspondence between the General Manager of Akros Chemicals and the competition law co-ordinator of Akzo Nobel.<sup>10</sup> The EC could not reach a definitive conclusion during the raid in relation to the privilege claimed over the first set of documents, which were subsequently placed in a sealed envelope. The EC official responsible for the investigation took the view that the second set of documents were not protected by privilege and had already taken copies and added them to the file, without placing them in a separate sealed envelope.

The question was referred to the CFI who in October 2003, prevented the EC (in an interim order) from reading the first set of documents pending the full proceedings in order to prevent serious and irreparable harm to the company. The CFI held the protection of professional privilege may also extend to written communications with a lawyer employed by an undertaking on a permanent basis. Interim relief was only granted for the first set of documents as the second set of documents were not in a sealed envelope and had already been placed on the EC's file, thus not satisfying the condition relating to urgency.

In September 2004, the ECJ overturned the decision of the CFI on the grounds the condition of "urgency" of such measures had not been satisfied. As the EC had already provisionally reviewed some of the documentation during the investigation, any harm which may result from a more detailed review of the documents was not sufficient to establish serious and irreparable harm to the company's interest.<sup>11</sup>

The EC considered that the President of the CFI applied a concept of professional privilege which has not been recognised and is not consistent with existing case law, in particular with *AM&S*. The EC also gave an undertaking that it would not allow third parties to have access to the first set of documents until judgment is given on the main application.

The decision by the ECJ in *Akzo Nobel* is a concern for companies involved in EC investigations. Although the EC might be prevented from explicitly relying on a document ultimately held to be privileged, its ability to review such a document during the course of its investigation has the potential to damage a party's interests.

In a positive move for companies involved in EC investigations, on 26 February 2007, the International Bar Association ("IBA") was granted leave by the CFI to intervene in the *Akzo Nobel* proceedings (Joined cases T-125/03R and T-253/03R). The CFI accepted the IBA's submissions that the case raised fundamental issues regarding the principle of legal professional privilege and held the judgment to be given in the full proceedings may significantly affect both the functioning of the

<sup>9</sup> *Akzo Nobel and Akros Chemicals (Akzo)*: Joined cases T-125/03R and T-253/03R.

<sup>10</sup> Akzo Nobel's competition law co-ordinator was a registered attorney in the Netherlands, a member of the Legal Department and employed by that undertaking on a permanent basis.

<sup>11</sup> *Akzo Nobel Chemicals & Akros Chemicals v Commission* (C7/04).

sector concerned and the interests of the IBA's members. However, as its application for leave to intervene was submitted after the expiry of the six-week period prescribed in Article 115(1) of the Rules of Procedure of the CFI, the IBA's intervention is limited to the submission of observations during the oral procedure.

The full proceedings in this case are still pending, where the issue of privilege will be raised again. Until such time, the position taken by the ECJ in *AM&S* remains.

### 3. The Modernisation Regulation

*Regulation 17 of the EC Treaty* as considered in *AM&S* has since been replaced by *Regulation 1/2003*, or the "Modernisation Regulation". The Modernisation Regulation has changed the implementation of *Article 81 of the EC Treaty*. While it does not affect the position of legal privilege under EC law, it does place a greater responsibility on individual national competition agencies and provides for the sharing of information between EU member states.

The Modernisation Regulation in fact permits the EC to access information from National Competition Authorities ("NCA"), through the European Competition Network ("ECN").<sup>12</sup> The ECN has the ability to exchange information among authorities across borders which can result in them exchanging legally privileged documents that were retrieved in a jurisdiction that does not provide privilege for an in-house lawyer, and pass that documentation to other jurisdictions that would normally not be able to view such communications.

It is important to note the distinction between the law applicable to the collection of information and the law applicable to the use of information in this context. In relation to the collection of information, under *Article 22* of the Modernisation Regulation, when an NCA conducts an investigation on its own territory on another NCA's behalf or at the EC's request, it does so under its national law.

In relation to the use of information, under *Article 12(1)* of the Modernisation Regulation, it is possible that the information collected can be received and used by the EC and the NCA even if it has been collected under rules that offer less protection to the client-lawyer communication than those of the EC or the receiving NCA. That is, the information may be used by the EC and the receiving NCA even though it could not have collected the information itself.

Notwithstanding the provisions of the Modernisation Regulation, the EC (and other regulatory authorities particularly in Financial Services) will in many cases propose a waiver of privilege and confidentiality in relation to documents and legal advice during the course of an investigation. There is no equivalent document in the EU to the US "Thompson Memorandum" and the subsequent "McNulty Memorandum".

It's therefore advisable that companies dealing within the EU must consider carefully that any communications may be subject to a wide variety of national laws and interpretations on privilege which were not contemplated when the communications came into existence.

## IV. PRACTICAL IMPLICATIONS: MANAGING THE RISKS AND CROSS-BORDER DISPUTES

### 1. EC Investigations: What is Privileged

- (a) Generally, the following communications are privileged in the context of an EC investigation:

<sup>12</sup> Together, the National Competition Authorities ("NCA") and the EC form a network of public authorities that act in the public interest and cooperate closely in order to protect competition. The network is called the European Competition Network ("ECN"). The ECN seeks to facilitate close cooperation between national competition authorities and the EC to ensure consistent application of Article 81 and 82 of the EC Treaty.

- Between the company and an external lawyer (subject to conditions in *AM&S*).
  - Preparation of internal reports by an in-house lawyer or other employee of advice received by an external lawyer (subject to conditions in *Hilti*).
- (b) Generally, the following communications are not privileged in the context of an EC investigation:
- Between the company and an in-house lawyer.
  - File notes prepared by in-house lawyers or other employees.
  - Requests from the company or an in-house lawyer to professional advisors for information or advice, other than an external lawyer qualified to practise in the EEA.
  - Preparation of internal reports by the company or an in-house lawyer of advice received by professional advisors, other than an external lawyer qualified to practise in the EEA.

## 2. Best Practice: Disclosure

- Understand local rules and disclosure requirements.
- Understand the consequences of transferring documents from a Civil Code country to a Common Law country - even on the basis of evaluation and review.
- Transferring documents from a Civil Code subsidiary to a Common Law subsidiary has a potential impact of disclosure.
- The reverse considerations apply: Common Law to Civil Code jurisdictions: disclosure may assist in making specific applications.

## 3. Best Practice: Privilege and Confidentiality

- Recognise the local/national law rules when corresponding across countries.
- Do not rely on informality: establish confidentiality wherever possible, particularly in dispute scenarios.

# THE PRACTICAL IMPLICATIONS OF PROPOSED RULE 502

*Ashish Prasad and Vazantha Meyers  
Mayer, Brown, Rowe and Maw LLP  
Chicago, IL*

The attorney-client privilege and work product doctrine are designed to protect the “full and frank” communications between attorneys and their clients necessary to give sound and reasoned legal advice.<sup>1</sup> One of the more difficult issues facing counsel today is preventing inadvertent disclosure of documents that are possibly protected by the attorney-client privilege or work product doctrine. Given the volume, nature and variety of electronically stored information (“ESI”) in large organizations today, litigation now often involves the review and production of thousands of documents, or more, and is often a time-consuming, burdensome and expensive process.<sup>2</sup> Counsel and their clients are often tasked with preventing the disclosure of privileged documents and a waiver of privilege in situations where, due to the large volume of ESI being produced, there is a substantial possibility of inadvertent production of privileged information. This article explores issues surrounding Proposed Rule 502 and the potential impact the enactment of the Rule will have in situations where privileged information is inadvertently disclosed.

## I. INADVERTENT DISCLOSURE

Because of the volume of documents involved, even the most diligent of privilege reviews of ESI could lead to the inadvertent disclosure of privileged documents. Although some courts will find a waiver only if the disclosure was intentional,<sup>3</sup> other courts are not sympathetic and will rule that a party that has mistakenly produced a privileged document has waived the privilege, despite efforts to prevent the disclosure of protected documents.<sup>4</sup> In *In re Sealed Case*, a company mistakenly disclosed a privileged document to the Internal Revenue Service. Counsel for the company argued that the attorney-client privilege had not been waived because the disclosure was involuntary and a result of a “bureaucratic error.”<sup>5</sup> The Circuit court held that the attorney-client privilege had been waived and refused to “distinguish between the various degrees of ‘voluntariness’” in waivers of attorney-client privilege.<sup>6</sup>

The majority of courts will find a waiver only if the disclosing party acted carelessly in disclosing the protected materials and failed to request its return in a timely manner. A multi-factor test has been developed to determine whether the inadvertent disclosure constitutes waiver of privilege protections.<sup>7</sup> For example, in *Buntin v. Becker*, the court considered the following factors:

1 *Upjohn Co. v. United States*, 449 U.S. 383, 389-93 (1981) (explaining the importance of the attorney-client privilege); *Hickman v. Taylor*, 329 U.S. 495, 500 (1947) (explaining the need to protect attorney work product). See Melissa A. Nunez, Note, *The Attorney, the Client, and...the Government?: A New Dimension to the Attorney-Client Privilege and Work Product Protection in the Post-Enron Era*, 82 Notre Dame L. Rev. 1311 (2007) for a discussion on the attorney-client and work product protections.

2 See, e.g., *Hopson v. City of Baltimore*, 232 F.R.D. 228, 244 (D.Md. 2005) (stating that electronic discovery may encompass millions of documents); *Rowe Entertainment, Inc. v. William Morris Agency*, 205 F.R.D. 421, 425-26 (S.D.N.Y. 2002) (stating that the cost of pre-production review for privileged and work product material could cost one defendant \$120,000 and another \$247,000, and that such review could take months).

3 See, e.g., *O'Mary v. Mitsubishi Electronics America, Inc.*, 59 Cal.App.4th 563, 69 Cal.Rptr.2d 389 (Cal. Ct. App. 1997); *People v. Murry*, 305 Ill. App.3d 311, 711 N.E.2d 1230 (Ill. App. Ct. 1999); *Premiere Digital v. Central Telephone*, 306 F. Supp.2d 1168 (D. Nev. 2005); *Trilogy Communications, Inc. v. Excom Realty, Inc.*, 279 N.J. Super. 442, 652 A.2d 1273 (N.J. Sup. Ct. Law. Div. 1994). See also *Hopson v. City of Baltimore*, 232 F.R.D. 228, 244 (D.Md. 2005) (discussing the three different approaches to analyzing inadvertent disclosure and privilege waiver cases); Kenneth S. Broun & Daniel J. Capra, *Getting Control of Waiver of Privilege in the Federal Courts: A Proposal For a Federal Rule of Evidence 502*, 58 SCLR 211(2006); Richard J. Healey, *Return to Sender? Inadvertent Disclosure of Privileged Information*, 28 Am. J. Trial Advoc. 615 (2005).

4 See, e.g., *In re Sealed Case*, 877 F.2d 976, 980 (D.C. Cir. 1989).

5 *Id.*

6 *Id.*

7 See, e.g., *Zapata v. IBP, Inc.*, 175 (F.R.D. 574, 576-577 (D. Kan. 1997); *Hydraflow, Inc. v. Enidine, Inc.*, 145 F.R.D. 626, 637 (W.D.N.Y. 1993); *State Compensation Insurance Fund v. Telanoff*, 70 Cal.App.4th, 82 Cal. Rptr.2d 799 (Cal. Ct. App. 1999); *Dalen v. Ozite Corp.*, 230 Ill. App.3d 18, 594 N.E.2d 1365 (Ill. App. Ct. 1992); *Baliwa v. State Farm Mutual*, 275 A.D.2d 1030, 713 N.Y.2d 376 (4th Dept. 2000).

(1) the reasonableness of the precautions to prevent inadvertent disclosure; (2) the time taken to rectify the error; (3) the scope of discovery; (4) the extent of the disclosure; and (5) the overriding issue of fairness.<sup>8</sup>

Parties sometimes address the risk of inadvertent disclosure of privileged materials and the burden of reviewing large volumes of documents early in the discovery process by entering into “clawback” or “quickpeek” agreements. The recently amended Federal Rules of Civil Procedure provide that if parties can enter into an agreement akin to a clawback or quickpeek agreement, the court may enter a case-management order adopting the agreement.<sup>9</sup> Pursuant to clawback agreements, the production of privileged data is presumed inadvertent and will not constitute a waiver; the party who receives the privileged data must return the privileged material when demanded until the privilege issue is resolved. Under a quickpeek agreement, the requesting party is allowed to review potentially responsive data and selects that data it wishes produced (the production of privileged data is presumed inadvertent and will not constitute a waiver). The producing party reviews for privilege only the selected document set and produces the non-privileged data from that set.

When effective, these agreements can limit the burdens and costs of privilege review in cases involving voluminous documents. However, while these agreements may be incorporated into court orders, the agreements and orders may not provide protection against arguments made by subsequent litigants not party to the original agreements that the privilege protections were waived when the materials were previously disclosed. In the absence of litigation or authority giving these agreements preclusive affect in subsequent litigation, these agreements will not ordinarily bind persons or entities not party to them. Further, while clawback and quickpeek agreements may be effective in the jurisdiction where the litigation is occurring, it may not protect that information in another jurisdiction or forum.

## II. SCOPE OF WAIVER

The inadvertent disclosure of protected material may result, under some circumstances, in the waiver of privilege with regard to all communications regarding the same subject matter. Some courts hold that when there has been a waiver of privilege as a result of an inadvertent disclosure, there is no subject matter waiver.<sup>10</sup> However, other courts have found that an inadvertent disclosure results in a subject matter privilege waiver.<sup>11</sup> In *In re Sealed Case*, for example, the court adopted a strict approach and found that an inadvertent disclosure resulted in not only the waiver of privilege as to the materials in question, but also as to the entire subject matter.<sup>12</sup> The court stated that a client must treat “the confidentiality of attorney-client communications like jewels – if not crown jewels.”<sup>13</sup> The court reasoned that the care that the company used in protecting its privileged communications reflected the importance of the confidentiality to the holder. The court further reasoned that its ruling that inadvertent disclosures resulted in waiver of privilege would encourage greater care in how corporations review materials for privilege and restraint in how freely corporations label documents privileged.<sup>14</sup>

Other courts have adopted a fairness approach to determining whether a disclosure results in subject matter waiver, which analyzes whether the waiver of privilege would be fair to the opposing party.<sup>15</sup> These courts recognize that withholding privileged materials limits the availability of information relevant to the litigation and useful to an opposing party. This recognition sometimes leads to a finding that subject matter privilege has been waived, especially when a party acts unfairly toward another party, or strategically or selectively discloses privileged communications.<sup>16</sup> In *Alpex*

8 727 N.E.2d 734 (Ind. Ct. App. 2000).

9 Fed. R. Civ. P. 26(f). The Advisory Committee’s note following the amendment to Rule 26(f) provides that parties “may attempt to minimize [the] costs and delays [associated with privilege review] by agreeing to protocols that minimize the risk of waiver.”

10 *Golden Valley Microwave Foods, Inc. v. Weaver Popcorn Co.*, 132 F.R.D. 204 (N.D. Ind. 1990); *International Digital Systems Corp., v. Digital Equipment Corp.*, 120 F.R.D. 445 (D. Mass 1988) (finding that there had been no subject matter waiver and that the inadvertent disclosure operates as “a waiver of the attorney-client privilege as to any documents disclosed by ‘inadvertence’”).

11 *In re Sealed Case*, 877 F.2d 976 (D.C. Cir. 1989).

12 *Id.*

13 *Id.* at 980.

14 *Id.*

15 See, e.g., *In re Grand Jury Subpoena*, 341 F.3d 331, 336–37 (4th Cir. 2003). See also, Broun, *supra* note 3, at 225; Jennifer A. Hardgrove, Note, *Scope of Waiver of Attorney-Client Privilege: Articulating a Standard That Will Afford Guidance to Courts*, 1998 U. Ill. L. Rev. 643 (1998).

16 See Broun, *supra* note 3, at 225–227 (citing *Parkway Gallery Furniture, Inc. v. Kittinger/Pennsylvania House Group, Inc.*, 116 F.R.D. 46 (M.D.N.C. 1987).

*Computer v. Nintendo Co., Ltd.*, the court allowed discovery of privileged documents after a party presented testimony related to those documents.<sup>17</sup> The court stated the aim of the subject matter waiver doctrine was to “prevent prejudice to a party and distortion of the judicial process that may be caused by the privilege-holder’s selective disclosure during litigation of otherwise privileged information,” and that subject matter waiver is found “only in the event that the initial disclosure is made during the course of litigation and results in prejudice to the opposing party.”<sup>18</sup>

### III. SELECTIVE WAIVER

In addition to the problems surrounding inadvertent disclosure of protected information, government entities, such as the U.S. Securities and Exchange Commission (“SEC”) and the Department of Justice (“DOJ”), are increasingly requesting that corporations turn over privileged documents to show cooperation with government investigations, waiving any privilege that may exist as to those documents.<sup>19</sup> Courts are not in agreement regarding whether that disclosure to a government agency or regulatory entity constitutes a general waiver of the privilege as to the information disclosed or a “selective waiver.”<sup>20</sup> Most courts will not find that a selective waiver has occurred, but will hold that the waiver of privileged or protected information to a government office or agency is a waiver for all purposes and all parties.<sup>21</sup> Some courts will find that a selective waiver occurred if the disclosure was made pursuant to an agreement of confidentiality with the government agency or regulatory entity.<sup>22</sup> A few courts have held that there is not a general waiver of privilege protection even if the disclosures to the government agency or regulatory entity were made in the absence of a confidentiality agreement.<sup>23</sup>

### IV. PURPOSE OF PROPOSED RULE 502

On January 23, 2006, the Chair of the U.S. House of Representatives Committee on the Judiciary, Honorable F. James Sensenbrenner, Jr., wrote to the Director of the Administrative Office of the U.S. Courts, requesting that the Judicial Conference begin a “rule-making on forfeiture of privileges.” Congressman Sensenbrenner explained that he was “informed that an absence of clarity on this subject, particularly as it pertains to the attorney-client privilege, is causing significant disruption and cost to the litigation process.” He went on to urge the Judicial Conference to adopt a rule that would “protect against the forfeiture of privilege where a disclosure is the result of an innocent mistake; permit parties and courts, to protect against the consequences of waiver by permitting disclosures of privileged information between the parties to a litigation; and allow persons and entities to cooperate with the government agencies by turning over privileged information without waiving all privileges as to other parties in subsequent litigation.”<sup>24</sup>

In March of 2006 the Advisory Committee on Evidence Rules of Practice and Procedure (the “Advisory Committee”) drafted proposed Rule 502 which was the subject of hearings at Fordham University on April 24 and 25, 2006. A revised version of Proposed Rule 502 was approved at the Advisory Committee’s June 22-23, 2006 meeting and published for public comment. In proposing Rule 502, the Advisory Committee recognized that there was a need for clear and uniform federal rules regarding subject matter waiver, inadvertent disclosures and the authority of the courts and parties to enter orders or agreements on privilege waiver.<sup>25</sup> Thus, Proposed Rule 502 attempts to address the issues that arise during litigation and a government investigation regarding attorney-client

17 1994 WL 330381 (S.D.N.Y. 1994).

18 *Id.* at 3.

19 Memorandum from Paul J. McNulty, Deputy Attorney General, to Heads of Department Components and United States Attorneys (Dec. 12, 2006) available at [http://www.usdoj.gov/dag/speech/2006/mcnulty\\_memo.pdf](http://www.usdoj.gov/dag/speech/2006/mcnulty_memo.pdf); Memorandum from Larry D. Thompson, Deputy Attorney General, to Heads of Department Components and United States Attorneys (Jan. 20, 2003), available at [http://www.usdoj.gov/dag/cltf/corporate\\_guidelines.htm](http://www.usdoj.gov/dag/cltf/corporate_guidelines.htm).

20 Braun, *supra* note 3, at 229-247 (discussing different approaches to analyzing selective waiver).

21 See, e.g., *Westinghouse Electronic Corp. v. Republic of the Philippines*, 951 F.2d 1414 (3d Cir. 1991); *McKesson HBOC, Inc. v. Superior Court of San Francisco*, 9 Cal.Reptr.3d 812 (Cal.Ct.App. 2004). See also, Mitchell, Preserving the Privilege: *Codification of Selective Waiver and the Limits of Federal Power over State Courts*, 86 B.U. L.Rev. 691 (2006) (stating that most states have rejected selective waiver).

22 See, e.g., *Teachers Insurance & Annuity Association of America v. Shamrock Broadcasting Co.*, 521 F.Supp. 638 (S.D.N.Y. 1981).

23 See, e.g., *Diversified Industries, Inc. v. Meredith*, 527 F.2d 596 (8th Cir. 1977).

24 Letter from Hon. F. James Sensenbrenner, Jr., Chair of the U.S. Cong. Comm. on the Judiciary to Leonidas Ralph Mecham, Dir. of the Admin. Office of the U.S. Courts 1 (Jan. 23, 2006 (on file with authors)).

25 Committee Note (1) – “It resolves longstanding disputes in the courts about the effect of certain disclosures of material protected by the attorney-client privilege or the work product doctrine – specifically disputes involving inadvertent and selective waiver.” (2) – It responds to the widespread complaint that litigation costs for review and protection of material that is privilege or work product have become prohibitive due to the concern that disclosure of protected information in the course of discovery (however innocent or minimal) will operate as a subject matter waiver of all protected information. This concern is especially troubling in cases involving electronic discovery.”



privilege and work product protection and provide a standard and uniform set of rules governing privilege waiver. Hearings on Proposed Rule 502 were held in January 2007 and the public comment deadline was February 15, 2007.

## V. PROPOSED RULE 502 AS IT WAS PRESENTED FOR PUBLIC COMMENT

The text of Proposed Rule 502 that was submitted for public comment provided as follows:

### **Rule 502. Attorney-Client Privilege and Work Product; Limitations on Waiver**

**(a) Scope of waiver.** – In federal proceedings, the waiver by disclosure of an attorney-client privilege or work product protection extends to an undisclosed communication or information concerning the same subject matter only if that undisclosed communication or information ought in fairness to be considered with the disclosed communication or information.

**(b) Inadvertent disclosure.** – A disclosure of a communication or information covered by the attorney-client privilege or work product protection does not operate as a waiver in a state or federal proceeding if the disclosure is inadvertent and is made in connection with federal litigation or federal administrative proceedings – and if the holder of the privilege or work product protection took reasonable precautions to prevent disclosure and took reasonably prompt measures, once the holder knew or should have known of the disclosure, to rectify the error, including (if applicable) following the procedures in Fed. R. Civ. P. 26(b)(5)(B).

**[( c ) Selective waiver.** – In a federal or state proceeding, a disclosure of a communication or information covered by the attorney-client privilege or work product protection – when made to a federal public office or agency in the exercise of its regulatory, investigative, or enforcement authority – does not operate as a waiver of the privilege or protection in favor of non-governmental persons or entities. The effect of disclosure to a state or local government agency, with respect to non-governmental persons or entities, is governed by applicable state law. Nothing in this rule limits or expands the authority of a government agency to disclose communications or information to other government agencies or as otherwise authorized or required by law.]

**(d) Controlling effect of court orders.** – A federal court order that the attorney-client privilege or work product protection is not waived as a result of disclosure in connection with the litigation pending before the court governs all persons or entities in all state or federal proceedings, whether or not they were parties to the matter before the court, if the order incorporates the agreement of the parties before the court.

**(e) Controlling effect of party agreements.** – An agreement on the effect of disclosure of a communication or information covered by the attorney-client privilege or work product protection is binding on the parties to the agreement, but not on other parties unless the agreement is incorporated into a court order.

**(f) Included privilege and protection.** – As used in this rule:

- (1) “attorney-client privilege” means the protection provided for confidential attorney-client communications, under applicable law; and
- (2) “work product protection” means the protection for materials prepared in anticipation of litigation or for trial, under applicable law.

## VI. PRACTICAL IMPLICATIONS OF PROPOSED RULE 502 AS IT WAS PRESENTED FOR PUBLIC COMMENT

While it is undeniable that the burdens placed on litigators and corporate clients to protect against disclosure of attorney-client privileged and work product information has increased as a result

of the growing volume of ESI, and it seems clear that Proposed Rule 502 as it was presented for public comment would reduce that burden to some extent, counsel for large organizations should be aware of the following issues related to the impact of Proposed Rule 502 as it was presented for public comment. Below, in the section of this paper entitled “Changes to Proposed Rule 502 Based on Public Comments,” we discuss how the Advisory Committee addressed certain of these issues.

1. Proposed Rule 502 as it was presented for public comment likely would have had no impact on state court actions where disclosure is initially made at the federal level. As illustrated above, state courts are split on how they analyze inadvertent disclosures, quick peek and clawback agreements, selective waiver and subject matter waiver. Because defendants cannot control the forum in which they are sued, in order to guard against possible privilege waiver in state court, litigants and corporate clients may not be able to effectively reduce the efforts expended on the review for privileged materials even if the initial disclosure takes place in federal court. Further, because the same set of documents are often reviewed for both federal and state proceedings simultaneously, the availability of the relief provided by Proposed Rule 502 as it was presented for public comment may not have reduced the burdens associated with reviewing for privilege, because state law considerations will drive the privilege review.
2. Proposed Rule 502(b) as it was presented for public comment attempted to reconcile the different approaches to analyzing inadvertent disclosures, but did not offer much guidance as to what constitutes “reasonable precautions” to prevent disclosures or “reasonably prompt measures” to rectify disclosures. The Committee Notes for Subdivision (b) referred to the split amongst the courts in dealing with inadvertent disclosures, and indicated that the rule opts for the middle-ground approach. The Committee Notes stated, “inadvertent disclosure of privileged or protected information in connection with a federal proceeding constitutes waiver only if the party did not take “reasonable precautions to prevent disclosure and did not make reasonable and prompt efforts to rectify the error.” Proposed Rule 502 as it was presented for public comment seemingly adopted the majority approach for analyzing whether an inadvertent disclosure results in a waiver of privilege protections. The rule effectively put litigants on notice that they must take substantial steps to prevent disclosures of privileged documents, but that they should have some protection if the disclosure was not intentional or the result of negligence.
3. Proposed Rule 502(a) as it was presented for public comment limited the application of subject matter waiver to instances where the “undisclosed communication or information ought in fairness ...be considered with the disclosed communication or information.” This language could have created contentious issues on when subject matter waiver has occurred. As it stands now, the courts are split on how to analyze subject matter waiver. Some courts lean toward finding subject matter waiver when there has been inadvertent disclosure and others hold that, only if the disclosure was intentional and fairness requires, will there be a subject matter waiver. The rule, as it was written, provided little guidance on distinguishing between an innocent disclosure and one that is done intentionally in order to gain a strategic advantage over the opposing party.

## VII. CHANGES TO PROPOSED RULE 502 BASED ON PUBLIC COMMENTS

The Advisory Committee met on April 12 and 13, 2007 to discuss proposed Rule 502 as it was submitted for public comment and consider all of the comments received. On May 15, 2007, the Advisory Committee issued a Report to the Standing Committee regarding the April 2007 meetings (“the Report”). The Advisory Committee made a number of modifications to the proposed Rule in its Report.

In the Report, the Advisory Committee summarized the changes to Proposed Rule 502 as it was presented for public comment, as follows:

1. Changes were made by the Style Subcommittee of the Standing Committee, both to the text as issued for public comment, and to the changes to the rule made at the April 2007 Evidence Rules Committee Hearing.

2. The text was clarified to indicate that the protections of Rule 502 apply in all cases in federal court, including cases in which state law provides the rule of decision.
3. The text was clarified to stress that Rule 502 applies in state court with respect to the consequences of disclosure previously made at the federal level – despite any indication to the contrary that might be found in the language of Rules 101 and 1101.
4. Language was added to emphasize that a subject matter waiver cannot be found unless the waiver is intentional – so that an inadvertent disclosure can never constitute a subject matter waiver.
5. The Committee relaxed the requirements necessary to obtain protection against waiver from inadvertent disclosure. As amended, the inadvertent disclosure provision assures that parties are not required to take extraordinary efforts to prevent disclosure of privilege and work product; nor are parties required to conduct a post-production review to determine whether any protected information has been mistakenly disclosed.
6. The protections against waiver by mistaken disclosure were extended to disclosures made to federal offices or agencies, on the ground that production in this context can involve the same costs of pre-production privilege review as in litigation.
7. The selective waiver provision – on which the Evidence Rules Committee had never voted affirmatively – was dropped from the Proposed Rule 502. The Evidence Rules Committee approved a separate report to Congress on selective waiver, setting forth the arguments both in favor and against the doctrine, and explaining the Committee's decision to take no position on the merits of selective waiver. The Evidence Rules Committee also prepared language for a statute on selective waiver to accompany that separate report to Congress; while the Committee took no position on the merits, it determined that the language could be useful to Congress should it decide to proceed with a separate selective waiver provision.
8. The Committee deleted the language conditioning enforceability of federal court confidentiality orders on agreement of the parties. It concluded that a federal order finding that disclosure is not a waiver should be enforceable in any subsequent proceeding, regardless of party agreement.
9. The definition of work product was expanded to include intangible information, as the work product protection under federal common law extends to all materials prepared in anticipation of litigation, including intangibles.

The text of the post-public comment Proposed Rule 502 is as follows:

#### VIII. RULE 502. ATTORNEY-CLIENT PRIVILEGE AND WORK PRODUCT; LIMITATIONS ON WAIVER

The following provisions apply, in the circumstances set out, to disclosure of a communication or information covered by the attorney-client privilege or work product protection.

- (a) Disclosure made in federal proceeding or to a federal office or agency; scope of waiver. – When the disclosure is made in a federal proceeding or to a federal office or agency and waives the attorney-client privilege or work-product protection, the waiver extends to an undisclosed communication or information in a federal or state proceeding only if:
- (1) the waiver is intentional;
  - (2) the disclosed and undisclosed communications or information concern the same subject matter; and
  - (3) they ought in fairness to be considered together.
- (b) Inadvertent disclosure. -- When made in a federal proceeding or to a federal office or agency, the disclosure does not operate as a waiver in a federal or state proceeding if:

- (1) the disclosure is inadvertent;  
 (2) the holder of the privilege or protection took reasonable steps to prevent disclosure; and  
 (3) the holder promptly took reasonable steps to rectify the error, including (if applicable) following Fed. R. Civ. P. 26(b)(5)(B).
- (c) Disclosure made in a state proceeding. – When the disclosure is made in a state proceeding and is not the subject of a state-court order, the disclosure does not operate as a waiver in a federal proceeding if the disclosure:  
 (1) would not be a waiver under this rule if it had been made in a federal proceeding; or  
 (2) is not a waiver under the law of the state where the disclosure occurred.
- (d) Controlling effect of court order. – A federal court may order that the privilege or protection is not waived by disclosure connected with the litigation pending before the court. The order binds all persons and entities in all federal or state proceedings, whether or not they were parties to the litigation.
- (e) Controlling effect of party agreements. – An agreement on the effect of disclosure is binding on the parties to the agreement, but not on other parties unless it is incorporated into a court order.
- (f) Controlling effect of this rule. – Notwithstanding Rules 101 and 1101, this rule applies to state proceedings in the circumstances set out in the rule. And notwithstanding Rule 501, this rule applies even if state law provides the rule of decision.
- (g) Definitions. – In this rule:  
 (1) “attorney-client privilege” means the protection that applicable law provides for confidential attorney-client communications; and  
 (2) “work-product protection” means the protection that applicable law provides for tangible material (or its intangible equivalent) prepared in anticipation of litigation or for trial.

In the Report, the Advisory Committee recognized that several changes needed to be made to the proposed Rule in order for it to have the impact that was intended when Rule 502 was initially proposed. The Advisory Committee recognized that the lack of enforcement in the state courts could potentially create a problem for corporations attempting to manage discovery efforts during a federal court action, and agreed that Rule 502 “must apply in state court actions where the question considered by the state court is whether a disclosure previously made in federal court constitutes a waiver.”<sup>26</sup> The Committee also made changes to the language used to define the requirements necessary to obtain protection against waiver from inadvertent disclosure so that it would be clear that parties will not be required to do anything extraordinary in order to seek protection against waiver. Under the modified rule, a party will be able to seek protection if it acted reasonably to prevent disclosure and took reasonably prompt measures to rectify the error, including following Fed. R. Civ. P. 26(b)(5)(B). This modification provides more guidance than the rule as it was presented for public comment on what a party will be required to show in order to seek protection against waiver.

One of the more notable post-public comment changes made to the proposed rule was the emphasis that subject matter waiver cannot be found unless the waiver was intentional. Inadvertent disclosure can never constitute a subject matter waiver under the rule as presently formulated. Another modification that will have a significant impact is the provision that states that a federal court order finding that privilege has not been waived will bind all persons and entities in all state and federal proceedings. Corporations will have a heightened ability to rely, even in state court, on agreements, such as “clawback agreements,” which become a part of court orders, and the management of large-scale discovery should be less burdensome as a result.

Of particular note is the removal of the selective waiver provision from the pre-public comment Proposed Rule 502. The Advisory Committee received several comments expressing concern that the selective waiver provision would add to the “culture of waiver” that has been created

<sup>26</sup> *Id.*

by government agencies' and regulatory entities' belief that it is appropriate to expect investigated corporations to disclose protected information and waive privilege protections to provide cooperation with the government's investigation.<sup>27</sup> Privilege protections allow an attorney to engage in the discussions necessary to uncover wrongdoing in a corporation, advise a client when it has been accused of wrongdoing and establish policies that prevent misbehavior. The fear was that this culture of waiver will cause individuals within the corporate environment to limit their communications with counsel because of the fear that those communications will subsequently be disclosed to the government in an effort to appear cooperative. The Advisory Committee listened to the comments for and against the selective waiver provision and decided to remove the provision from Proposed Rule 502.

## IX. CONCLUSION

As noted earlier, a principal purpose behind the protections for attorney-client privilege and work product is to encourage full and frank communication between an attorney and his or her client. The growth in the volume, nature and complexity of ESI in recent years has increased the danger of privilege waivers resulting from innocent mistakes in the review and production of documents. Full and frank communication between an attorney and his or her client will be limited if there is no protection from waiver in these circumstances. By limiting the circumstances in which the inadvertent disclosure of privileged information can result in a waiver, and the scope of the waiver that could potentially be found, Proposed Rule 502 goes a long way toward mitigating this risk, and may be a welcome development for attorneys and their clients.

---

<sup>27</sup> The Decline of the Attorney Client Privilege in the Corporate Context: Survey Results (2006) at 3, *available at* <http://www.acca.com/Surveys/atryclient2.pdf>.

# THE SEDONA GUIDELINES: BEST PRACTICES ADDRESSING PROTECTIVE ORDERS, CONFIDENTIALITY & PUBLIC ACCESS IN CIVIL CASES

---

*A Project of The Sedona Conference® Working Group  
on Protective Orders, Confidentiality &  
Public Access (WG2)\**

*March 2007 Post-Public Comment Version*

*Editorial Team:*

Professor Laurie Dore  
Ronald J. Hedges  
Kenneth J. Withers

*Advisory Editorial Team:*

The Hon. J. Thomas Marten, District of Kansas, Wichita, KS  
Prof. Stephen Saltzburg, George Washington University, Washington, D.C.  
Prof. Jay Tidmarsh, Notre Dame Law School, Notre Dame, IN

### **Dedication to H. Brent McKnight (1952-2004)**

H. Brent McKnight, a member of the Steering Committee of this Working Group, passed away suddenly and unexpectedly in the year between the group's first meeting and the publication of its first draft Principles & Best Practices. Years before, Brent had been a Morehead Scholar who spent an internship working for Senator Sam Ervin, reading through 1,100 boxes of Watergate documents. As a Rhodes Scholar, studying religious philosophy under Bishop Stephen Neill at Oxford, Brent decided that justice was a true goal of religion and that justice would be his career. Prosecutor, state court judge, United States Magistrate Judge, United States District Judge, teacher, writer, scholar, husband, father, colleague and friend: through all this he infused his desire to see justice. And through all this, we were struck by the fact that he kept asking us to make sure he was getting it right. We were amazed at his humility wedded to intelligence; his sense of humor wedded to hard work. To you, Brent, we dedicate our work. May it help others advance the justice to which you devoted your life.

*Alan Blakley*

---

\* The opinions expressed herein are consensus views of the editors and authors, and do not necessarily represent the views of any individual participants, members, or authors or any of the organizations to which they belong or clients they represent, nor do they necessarily represent official views of The Sedona Conference®

## INTRODUCTION

We live in an open and democratic society that depends upon an informed citizenry and public participation in government. Open public meetings laws and federal and state freedom of information laws facilitate such participation by providing citizens with a right of access to information concerning their government. Indeed, the First Amendment protects “the stock of information from which members of the public may draw.” *Richmond Newspapers, Inc. v. Virginia*, 448 U.S. 555, 576(1980).

At the same time, our society places high premiums upon personal privacy, property rights, and individual autonomy. The United States Supreme Court has elevated some aspects of privacy to a constitutional (albeit rather amorphous) right. Our country has long valued entrepreneurial confidentiality as a key to social and material progress, promoting individual initiative, private enterprise, and technological innovation and the right to protect the property created by that enterprise. Courts have made clear in the civil litigation context that litigants may have privacy and proprietary rights in certain of the information produced during the discovery process. *Seattle Times v. Rhinehart*, 467 U.S. 20 (1984). Nevertheless, the primary responsibility to protect those rights at each stage of the litigation rests on the attorney for the producing party.

This inherent tension between public access to information about government activities and the desire to protect personal privacy, intellectual property and confidential business information comes to a head in the debate concerning confidentiality in litigation. As with the legislative and executive branches of government, our democratic society depends upon public participation in and access to the judicial process. Access to the courts both improves the operation of the judicial system and fuels the informed discussion essential to democracy. Public access to judicial proceedings facilitates public monitoring of our publicly-created, staffed, and subsidized judicial system. Fair and open judicial proceedings and decisions encourage public confidence in and respect for the courts – a trust essential to continued support of the judiciary. A public eye on the litigation process can enhance prompt, fair and accurate fact-finding and decision-making. Perjury is deterred, witnesses may step forward, and judgments may be tempered with greater care and deliberation. A public trial also educates citizens about the justice system itself as well as its workings in a particular case.

Unlike the legislative and executive branches of government, however, the primary function of the courts is the resolution of the “cases and controversies” before them, and public access to certain stages of civil lawsuits casts light beyond the judicial process itself. Indeed, civil litigation most often involves disputes between private parties who are drawn into the courts reluctantly or even involuntarily. These court proceedings may require the disclosure of intimate personal or financial information or the disclosure of trade secrets or confidential marketing, research, or commercial information may be at stake. Public access to the pretrial, trial or settlement stages of those cases thus might jeopardize legitimate privacy or proprietary interests of the litigants. Moreover, public access may hamstring the litigants’ ability to resolve their dispute in a mutually agreed manner.

Many threshold questions were addressed by WG2: Do litigants give up a measure of their privacy and autonomy when they enter the doors of the public courthouse in order to resolve their disputes? To what extent should court rules on protective orders, confidentiality and public access take into consideration the possibility that producing parties (or non-parties) did not voluntarily choose the dispute resolution forum? How is a court to honor the right of public access to judicial proceedings while protecting privacy and property interests? In which cases must such privacy and proprietary interests bow to a broader public interest?

Virtually every federal circuit court of appeals and state court of last resort that has spoken on these matters has determined that courts are required to weigh the public interest in the particular proceeding or stage of the litigation against the private interest in maintaining the confidentiality of the material under consideration.<sup>1</sup>

On the public interest side, the general rule announced by the United States Supreme Court is that the public's right of access to material produced in connection with a particular pretrial or trial proceeding arises when (1) the proceeding has historically been open and (2) public access plays a significant role in the proper functioning of the process. *Press-Enterprise Co. v. Superior Court*, 478 U.S. 1, 8 (1986); *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 605-06 (1982). In doing so, the Court and the case law relying on this general rule have made distinctions between, for instance, discovery as historically private exchanges conducted by the parties, and trials as historically open proceedings in which the public has an interest, directly or as a matter of public accountability. *Baxter International v. Abbott Labs*, 297 F.3d 544 (7th Cir. 2002).

On the private interest side, statutes and court rules have declared some types of proceedings (e.g., juvenile, child abuse, adoption, and guardianship) to be presumptively closed to the public, and certain types of information (e.g., the "personal identifiers" specified in proposed Fed. R. Civ. P. 5.2, [http://www.uscourts.gov/rules/Rules\\_Publication\\_August\\_2005.pdf#page=55](http://www.uscourts.gov/rules/Rules_Publication_August_2005.pdf#page=55), including Social Security numbers, dates of birth, financial account numbers and names of minor children) to be presumptively eligible for exclusion from part or all of the court record available to the public. Case law has also given great weight to the private interest in protecting bona fide trade secrets and confidential proprietary information, such as marketing plans, employee training manuals, computer source code, and customer lists. Less weight has been given to unsupported claims of confidentiality for broadly designated business information. *Citizens First National Bank of Princeton v. Cincinnati Insurance Co.*, 178 F.3d 943 (7th Cir. 1999).

Attorneys must regularly consider the differing levels of protection (or conversely, the levels of public access) that may be afforded to materials exchanged in the course of discovery, and materials filed with the court. Further, when considering materials to be filed, attorneys and courts must distinguish between the levels of protection (or conversely, the levels of public access) afforded to materials being filed in relation to a nondispositive matter and materials that relate to the merits of the case.

In the discovery context, there is no presumption of public access to unfiled discovery. *Seattle Times Co. v. Rhinehart*, 467 U.S. 20 (1984). There is also no prohibition against a party disseminating information obtained through discovery. Attorneys may seek to enforce their clients' privacy and proprietary interests by agreement or may seek a protective order under Fed. R. Civ. P. 26(c) by showing "good cause." In the context of filing material with a court, a threshold presumption of public access exists. If the material relates to a nondispositive matter, a "good cause" determination by the court in issuing an order to seal is sufficient to overcome the presumption of public access. *Chicago Tribune Company v. Bridgestone/Firestone, Inc.*, 263 F.3d 1304 (11th Cir. 2001). However, if the material relates to the merits of the case, an order to seal must be supported by a determination of compelling need that overcomes the presumption of public access. *Leucadia, Inc. v. Applied Extrusion Technologies, Inc.*, 998 F.2d 157 (3d Cir. 1993). Thus, materials "designated as confidential under a protective order... will lose confidential status (absent a showing of 'most compelling' reasons) if

1 "[T]he ordinary showing of good cause which is adequate to protect discovery material from disclosure cannot alone justify protecting such material after it has been introduced at trial. This dividing line may in some measure be an arbitrary one, but it accords with long-settled practice in this country separating the presumptively private phase of litigation from the presumptively public." *Poliquin v. Garden Way Inc.*, 989 F.2d 527, 533 (1st Cir. 1993). *Compare, Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 786-87 (3d Cir. 1994) (confidentiality order over unfiled settlement agreement); *Phillips v. General Motors Corp.*, 307 F.3d 1206, 1210-12 (9th Cir. 2002) (confidentiality order over unfiled settlement information) *with*, *In Re Providence Journal*, 293 F.3d 1, 9-11 (1st Cir. 2002) (access to trial exhibits); *Lugosch v. Pyramid Co.*, 435 F.3d 110, 119-20 (2d Cir. 2006) (access to summary judgment materials); *Hartford Courant Co. v. Pelligrino*, 380 F.3d 83, 90-96 (2d Cir. 2004) (access to dockets); *In Re Cendant Corp.*, 260 F.3d 183, 192-93 (3d Cir. 2001) (access to bids and bidding auction); *Leucadia v. Applied Extrusion Technologies*, 998 F.2d 157, 161-65 (3d Cir. 1993) (access to nondiscovery pretrial motions); *Virginia Dept. of State Police v. Washington Post*, 386 F.3d 567, 574-77 (4th Cir. 2004) (access to summary judgment materials); *Second v. Van Waezenberghe*, 990 F.2d 845, 848 (5th Cir. 1993) (access to transcript and final order of permanent injunction as part of settlement agreement); *Baxter Internat'l, Inc. v. Abbott Labs.*, 297 F.3d 544, 545-56 (7th Cir. 2002) (access to documents on appeal); *Grove Fresh Distrib., Inc. v. Everfresh Juice Co.*, 24 F.3d 983, 896-97 (7th Cir. 1994) (access to court files); *In Re Neal*, 461 F.3d 1048, 1052 (8th Cir. 2006) (access to creditor list in bankruptcy proceeding); *Kamakana v. City and County of Honolulu*, 447 F.3d 1172, 1178-81 (9th Cir. 2006) (access to summary judgment materials); *EEOC v. National Children's Center*, 98 F.3d 1406, 1409-10 (D.C. Cir. 1996) (access to consent decree).



introduced at trial or filed in connection with a motion for summary judgment.” Manual for Complex Litigation, §11.432 (4th Edition 2004).

The relative strength of the public’s interest in access versus a litigant’s interest in privacy, property or confidentiality evolves with the stage of litigation, so that what constitutes “good cause” to restrict public access during discovery or non-dispositive stages of the proceeding does not equate to the “compelling need” necessary to restrict public access at another stage. This complexity appears to have generated widespread confusion in practice. The reported case law, supported by comments received by the Working Group in response to the Revised April 2005 Public Comment Version of this publication, demonstrates that litigants frequently move to seal docket entries, court filings, or whole proceedings, citing standards applicable only in the discovery or non-dispositive context. Likewise, judges across the country are routinely presented with stipulated discovery protective orders that the parties claim govern filings on the merits. Under the pressure of court workloads, some judges may be tempted to improperly forgo the individual determinations necessary to seal court documents, and instead issue orders in accordance with the parties’ stipulations. *See, e.g., Citizens First National Bank of Princeton*, 178 F. 3d 943, 944 (7th Cir. 1999) (“Instead of [making a determination] he granted a virtual carte blanche to either party to seal whatever portions of the record the party wanted to seal. This delegation was improper.”). More recently, the process of modernizing and automating court filing and case management systems has revealed age-old informal practices under which court clerks or counsel themselves routinely sealed filings without any judicial determination whatsoever.

The electronic age and the requisite process of modernization has also led to a new concern. The conversion of presumptively public court records from paper-based filing systems accessible only at the courthouse itself, to electronic records potentially available via remote Intranet or Internet connections, has changed the analysis of the weight to be given to privacy concerns. While in the past, the likelihood that an individual or business would go to the effort of using court files to access personal information for private gain was remote, the automation of these records has made the harvesting of personal information for commercial use a viable, and indeed quite profitable, business. This has given rise to a new legal dialectic which recognizes a responsibility on the part of government to protect the confidentiality of personal information, and perhaps confidential business information, that it requires citizens to provide as part of the civil and criminal justice systems.

For these reasons, the Working Group determined that the bench and bar would benefit from suggestions for “best practices” regarding public access to court files and proceedings in civil litigation, together with illustrations reflecting common situations that litigants and judges are likely to face.

Chapter One deals with discovery, and it is placed first because it is clearly distinct from other aspects of litigation, in that it is largely private and party-controlled – until the fruits of discovery are filed with the court for consideration of the merits of the case. Chapter Two deals with the important administrative functions of the court-the procedures for filing, maintaining the docket, and handling court-generated documents such as opinions and judgments. Here the applicable standards for sealing are different from the standards for discovery protective orders, due to the greatly increased public interest in the workings of the court and the greatly narrowed focus of the materials involved to those that deal with the merits of the case. Chapter Three goes to the core judicial function with the greatest public interest-the trial itself, including jury selection and the evidence presented. Chapter Four takes up the question of settlements, which was not a primary focus of the Working Group, but has been a lightning rod for press coverage and legislative attention. Here the analysis is extended to consider settlements as private agreements between parties until the parties choose, or feel compelled, to invoke the supervisory or enforcement powers of the courts. Finally, Chapter Five explores the implications of the transformation of the courts from repositories of largely paper-based information to managers of digital information. In particular, Chapter Five examines the increased attention given to protecting personal information as courts redesign their processes and explore novel questions of public access in an electronic age.

## CHAPTER I DISCOVERY

### **Principle 1    There is no presumed right of the public to participate in the discovery process or to have access to the fruits of discovery that are not submitted to the court.**

The American federal civil litigation system is premised on the just, speedy and inexpensive resolution of disputes. Fed. R. Civ. P. 1. The scope of discovery under the Federal Rules of Civil Procedure is intended to be broad. Parties may obtain discovery regarding “any matter, not privileged, that is relevant to the claim or defense of any party,” and for good cause shown, may obtain broader discovery relevant to the subject matter of the dispute. The information requested and produced during the discovery phase of civil litigation “need not be admissible at trial if [it] appears reasonably calculated to lead to the discovery of admissible evidence.” Fed. R. Civ. P. 26(b)(1).

Unlike the civil law system in Europe and elsewhere, in the American civil litigation system the parties themselves develop the facts they need for trial through the discovery process outlined in the Federal Rules of Civil Procedure and state equivalents. These rules delegate to private parties the inquisitorial powers of the court, including the right to inspect and copy documents, the right to conduct depositions, and the right to compel non-parties to testify or produce documents. *Seattle Times Co. v. Rhinehart*, 467 U.S. 20 (1984) (“Thus, the [Federal Rules of Civil Procedure] often allow extensive intrusion into the affairs of both litigants and third parties.” 467 U.S. 20 at 30; “The Rules do not distinguish between public and private information.” 467 U.S. 20 at 36) The court does not usually involve itself in the conduct of civil discovery, although it enforces procedural rules and may be called upon to decide discovery disputes. Generally, the fruits of discovery (documents, answers to interrogatories, deposition testimony, etc.) are not filed unless these are being used as evidence, either at trial or in connection with a discovery dispute or other pretrial proceeding, or unless the court orders that these be filed. Fed. R. Civ. P. 5(d).<sup>2</sup>

Pretrial discovery that is simply exchanged between the parties is not a public component of a civil trial. *See, e.g., Seattle Times Co. v. Rhinehart*, 467 U.S. 20, 33 (1984) (“pretrial depositions and interrogatories are not public components of a civil trial . . . and, in general, they are conducted in private as a matter of modern practice”). There is thus no presumed right of public access to the discovery process or the fruits of discovery in the hands of a party. However, as discussed below, a party is not prohibited from voluntarily disclosing any information received during discovery unless the party has agreed otherwise or unless the court, upon a showing of good cause, enters a protective order pursuant to Fed. R. Civ. P. 26(c) or its state equivalents. A party’s ability to enter into such agreements, and the court’s ability to enter such orders, may be limited by statute or rule. *See, e.g., Fla. Stat. Ann. §69.081; Texas R. Civ. P. 76a.*

#### Best Practices

1. Attorneys should cooperate in efficiently exchanging information in civil litigation. Such cooperation includes an early, full discussion of the scope of discovery and the treatment of potentially discoverable materials that the parties deem confidential or private, to avoid later pretrial litigation of this issue.
2. A party may object to the discovery of otherwise relevant and non-privileged information it claims is confidential or private. Such an objection should be the basis for negotiation with the requesting party over the procedure for producing the requested discovery to protect legitimate privacy and confidentiality interests. If no agreement is reached, a party may apply to the court for a protective order under Fed. R. Civ. P. 26(c) or its state equivalents.

<sup>2</sup> It is sometimes argued that discovery requests are made to threaten the release of “sensitive” information and coerce settlement. This is a matter of professional responsibility within the scope of ethics rules and not addressed by the Working Group. *See* Model Rules of Professional Conduct (2001), Preamble; A Lawyer’s Responsibilities (5) (“A lawyer should use the law’s procedures only for legitimate purposes and not to harass or intimidate others”); Rule 3.4(d) (“A lawyer shall not . . . in pretrial procedure, make a frivolous discovery request. . .”)

3. If the parties agree to produce information under the terms of a stipulated protective order, the court should enter such order upon a showing of good cause, subject to a later determination of the confidentiality of specific documents in the event of a challenge to the confidentiality designation.

### Examples

1. The attorneys representing the parties in a class action race discrimination lawsuit against a large corporation confer pursuant to Fed. R. Civ. P. 26(f) and map out the discovery phase of the litigation. The lawsuit alleges discrimination in pay and promotions throughout the company. It is anticipated that the plaintiff will serve a broad discovery request seeking current and historical information regarding employee pay and promotions. It is also anticipated that the defendant will object to public disclosure of employee pay, citing employee morale and competitive interests. The attorneys negotiate a procedure for the production of the relevant information in bulk form, redacting any “personal identifiers” in the data, and enter into a confidentiality agreement or stipulate to a protective order that would permit large volumes of information to be reviewed and exchanged without compromising privacy and confidentiality interests.
2. Same facts as Example 1, but the attorneys were unable to reach agreement. The plaintiff serves its discovery request, as anticipated, and the defendant objects. Under the court’s rules, the attorneys must attempt to resolve discovery disputes before filing any motions. During the required meeting, the defendant flatly refuses to produce the requested data, and the plaintiff threatens to obtain the data from other sources and publish it on the Internet. The plaintiff then moves to compel discovery and the defendant counter-moves for a protective order. Three months and several hundred billable hours later, the court grants both motions in part, fashioning a protective order similar to that reached voluntarily in Example 1.

**Principle 2 Absent an agreement between the parties or a court order based on a showing of good cause, a litigant is not precluded from disclosing the fruits of discovery to non-parties.**

Absent an agreement between the parties or an order to the contrary, a party is free to share the fruits of discovery obtained during litigation with others who are not parties to the lawsuit. *Phillips v. General Motors Corp.*, 307 F.3d 1206, 1210 (9th Cir. 2002); *Harris v. Amoco Prod. Co.*, 768 F.2d 669, 683-84 (5th Cir. 1985).

In some cases, a producing party has legitimate reasons to limit the dissemination of certain information exchanged in the normal course of discovery. Because broad discovery is generally allowed, and given the nature of certain disputes in the civil justice system, the rules of discovery often require disclosure of private, confidential information involving matrimonial, financial, medical or family matters, or in commercial cases, trade secrets and other confidential business information. In order to facilitate the efficient exchange of information during discovery, parties may enter into agreements or stipulations designed to maintain the confidentiality of material produced during discovery.

In the absence of an agreement between the parties, a producing party has the right to object to the production of particular material on the basis of “annoyance, embarrassment, oppression, or undue burden,” and seek a protective order under Fed. R. Civ. P. 26(c). In appropriate cases, a party may seek an order “that a trade secret or other confidential research, development, or commercial information not be revealed or be revealed only in a designated way.” Fed. R. Civ. P. 26(c)(7). The court is required to make a finding of “good cause” to support a protective order.

To avoid a costly and time-consuming document-by-document determination of “good cause” by the court during the discovery, parties who anticipate that the scope of discovery will likely include private or confidential information often seek the court’s *imprimatur* under Fed. R. Civ. P. 26(c) on a stipulated agreement regarding confidentiality. These “umbrella” protective orders, most often found in large or complex cases, are frequently entered without judicial assessment of the

specific documents or information the disclosure of which will be limited by the protective order, although a more generalized finding of “good cause” is still required. Parties may demonstrate reliance on a protective order entered as a case management tool. However, such an order is insufficient justification by itself for the court to enter a sealing order if documents subject to the protective order are to be filed.

In determining whether good cause exists to issue or uphold a protective order under Fed. R. Civ. P. 26(c), a court is required to balance the parties’ asserted interest in privacy or confidentiality against the public interest in disclosure of information of legitimate public concern. *Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 786 (3d Cir. 1994). Judicial restraints placed only upon the disclosure and use of information exchanged in discovery do not restrict “a traditionally public source of information,” *Seattle Times Co. v. Rhinehart*, 467 U.S. at 33. “When directed solely at discovery materials, protective orders are not subject to the high level of scrutiny required by the Constitution to justify prior restraints; rather, courts have broad discretion at the discovery stage to decide when a protective order is appropriate and what degree of protection is required.” Manual for Complex Litigation, at 11.432, p. 66 (4th Edition, 2004) (citing *Seattle Times Co. v. Rhinehart*, 467 U.S. at 36-37). Therefore, given that the public shares the parties’ interest in a judicial system that can efficiently resolve disputes, the good cause standard generally should be considered to be satisfied if the parties can articulate a legitimate and particularized need for privacy or confidentiality, in those instances where the protective order will apply only to the disclosure of information exchanged during discovery. If a challenge is made, “good cause” must be shown based on the circumstances existing at the time of the challenge. *Leucadia, Inc. v. Applied Extrusion Techs., Inc.*, 998 F.2d 157 (3d Cir. 1993). Because of the limited scope and provisional nature of the umbrella protective order, the court need not conduct a detailed inquiry into the nature of the information at issue, which courts are sometimes unwilling or often practically unable to do, where much or all of the information at issue may not be used in connection with the determination of the merits of the dispute.

#### Best Practices

1. Attorneys should counsel their clients that, absent an agreement or order issued upon a showing of good cause, there is no restriction on dissemination of documents and other information exchanged during discovery. A party desiring restrictions on dissemination of the fruits of discovery should approach the opposing party as soon as it becomes apparent that some type of restriction on dissemination is necessary.
2. Attorneys should assess in each case whether a protective order restricting dissemination of information produced during discovery is necessary.
3. An umbrella protective order or confidentiality agreement should provide a procedure for confirmation or challenge of the confidentiality designations made as to particular documents, including timely notice to the producing party that the designation is being challenged to enable the producing party to promptly seek protection and to prove that the particular information qualifies for judicial protection.
4. Attorneys should avoid excessive and unjustified designation of documents as “confidential” under a protective order.

#### Examples

1. Using the race discrimination lawsuit example outlined under Principle 1 above, at the same time the attorney representing the corporate defendant notifies the attorney representing the plaintiff that she is willing to produce her client’s payroll and promotion database, she states that it contains private financial information on the employees of her client. She says she will only produce this database if the plaintiff’s counsel is agreeable to enter a protective order. The plaintiff’s attorney agrees and they negotiate a proposed protective order which:

- a. defines the information and documents it will protect, encompassing the type of information that could upon proof be the appropriate subject of a protective order;
- b. establishes a procedure whereby plaintiff's attorney may notify defendant's counsel if she believes that certain documents designated as "confidential" by the defendant should not be treated as such; and
- c. provides that within a specified period after being notified of plaintiff's counsel's objection to certain confidentiality designations, if defendant's counsel wishes to maintain the confidentiality of the challenged documents, the defendant's counsel must seek a protective order pursuant to Fed. R. Civ. P. 26(c). The protective order makes clear that if the confidentiality designation is challenged, the court is to make a *de novo* determination of whether there is good cause to restrict the dissemination of the challenged information.

**Principle 3 A protective order entered under Fed. R. Civ. P 26(c) to facilitate the exchange of discovery materials does not substitute for the individualized judicial determination necessary for sealing such material, if filed with the court on a non-discovery matter.**

Protective orders sometimes purport to do more than restrict the parties from sharing the fruits of discovery. They often include a provision allowing materials deemed "confidential" to be filed with the court under seal without any further order. Such an agreement between the parties may be appealing. Courts are understandably disinclined to interfere with a matter agreed upon by the parties, particularly considering the court's limited time and resources. However, no agreement between the parties should substitute for the individualized and particularized showing that must be made before any materials are filed under seal, at least for non-discovery purposes. Moreover, given the presumption of public access to filed materials, that showing must be made under the stricter standards described in Chapter 2 rather than the "good cause" standard of Fed. R. Civ. P. 26(c) used for the issuance of protective orders.

Although protective orders with "sealing" provisions appear to be common, federal circuit courts have questioned the enforceability of protective orders that serve to seal material filed with the court, primarily because such sealing implicates the public's qualified right of access to court records. For example, according to the Third Circuit, "[T]he burden of justifying the confidentiality of each and every document sought to be covered by a protective order remains on the party seeking the protective order; any other conclusion would turn Rule 26(c) on its head." *Leucadia, Inc. v. Applied Extrusion Techs., Inc.*, 998 F.2d 157, 166 (3d Cir. 1993). Similarly, the Sixth Circuit held that protective orders under Fed. R. Civ. P. 26(c) authorizing the sealing of documents that either party "considers ... to be of a confidential nature" is facially overbroad. *Procter & Gamble Co. v. Bankers Trust Co.*, 78 F.3d 219, 227 (6th Cir. 1996).

Particularly in cases with large quantities of material to be produced in discovery, a threshold showing of "good cause" over broad categories of material may be sufficient for the issuance of a protective order under Fed. R. Civ. P. 26(c). The purpose of the order would be to facilitate the cooperative exchange of voluminous discovery. Protective orders must not confuse the confidentiality of material produced in discovery with the filing of such materials under seal. That issue is discussed in detail in Chapter 2.

#### Best Practices

1. Attorneys should advise their clients that a protective order entered without an evidentiary showing only restricts the dissemination of designated documents and information so long as the need for confidentiality is not successfully challenged by another party or by an intervening third party, and so long as the information does not need to be filed in court or used as evidence at trial. The party wishing to prevent the dissemination of information may eventually be required to prove the basis for protecting specific information, even if not required to do so at the time the information is produced.

2. Protective orders which purport to cover both the exchange of documents in discovery and the filing of documents with the court would better conform with the legal standard if such protective orders were: (1) narrowly drafted; (2) kept the burden on the designating party to demonstrate good cause whenever the need for confidentiality is questioned; and (3) provided a procedure to establish a proper basis for sealing at the time the material is actually filed with the court for any purpose other than a discovery dispute.
3. Protective orders entered without evidentiary findings should provide a mechanism to establish whether information designated as confidential should be sealed if filed with the court. For instance, such an order could provide that a party “lodge” a protected document with the court pending a motion to seal from the designating party.<sup>3</sup> If the designating party files a motion to seal the court record within a reasonable period of time, a determination is then made as to whether the particular information should remain under seal. The fact that information was designated as confidential pursuant to the protective order is not dispositive in determining that the information should be sealed in connection with a determination on the merits.
4. As an alternative to the practice outlined above, the parties could agree in the protective order to provide reasonable notice to the designating party that it intends to file documents designated as confidential in court. The designating party should move within a reasonable period of time to have the specific documents sealed. Again, a judicial determination must be made as to whether the sealing of the particular records is warranted.

#### Examples

1. In the class action race discrimination lawsuit outlined under Principle 1 above and the protective order discussed as an example under Principle 2 above, the plaintiff’s counsel intends to file the payroll and promotion database which has been designated as confidential as an exhibit in opposition to a motion for summary judgment. Pursuant to the terms of the order, plaintiff’s counsel notifies defendant’s counsel in advance of the filing that the designation is challenged, triggering the defendant’s obligation to file a motion for a sealing order with respect to the challenged information. Alternatively, plaintiff’s counsel temporarily files the payroll and production database with the opposition papers, ensuring that the material is not made public and likewise triggering the defendant’s obligation.
2. Using the above example, plaintiff’s counsel also notifies defendant’s counsel that pages from one of defendant’s outdated employee handbooks will be filed in connection with a summary judgment motion. These pages had been designated as confidential pursuant to the protective order. Defendant’s counsel decides that it is not necessary to seek to have this outdated information sealed. Plaintiff’s counsel is permitted to file the information in open court, and the confidentiality designation with respect to that information is waived.

#### **Principle 4 On a proper showing, non-parties should be permitted to intervene to challenge a protective order that limits disclosure of discoverable information.**

A party involved in parallel or subsequent litigation should be permitted to present arguments that a protective order should be modified to allow it access to the allegedly confidential documents. A court deciding whether its protective order should be modified to allow a party to such litigation access to documents should consider the standards of relevance and efficiency articulated in Fed. R. Civ. P. 26, including considerations of annoyance, embarrassment, and oppression under Fed. R. Civ. P. 26(c). However, the public disclosure of information of a private or sensitive nature in one lawsuit should not necessarily subject a party to repeated disclosure of the same information in subsequent litigation, if there is good cause for protecting it from disclosure.

---

<sup>3</sup> A “lodged” document is a document submitted to the court in conjunction with the filing of a motion to allow the document to be filed under seal. The “lodged” document itself is not considered part of the filing. If the court denies the motion, the document is returned to the submitting party. See, e.g., N.D. Cal. Civ. L. R. 79-5(d). (“Sealed or Confidential Documents: Motion to File Under Seal”)

Most courts that have considered the question hold that the media, public interest groups, and other third-parties have standing to intervene in a civil case for the limited purposes of opposing or seeking modification or rescission of a protective order entered pursuant to Fed. Rule Civ. P. 26(c) when they assert that the public interest is served by disclosure. See *Grove Fresh Distributors v. Everfresh Juice Co.*, 24 F.3d 893, 896 (7th Cir. 1994); *Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 777 (3d Cir. 1994); *In re Alexander Grant & Co. Litigation*, 820 F.2d 352, 354-55 (11th Cir. 1987); *CBS v. Young*, 522 F.2d 234, 237-38 (6th Cir. 1975). Courts have found standing without any showing that persons subject to an order limiting disclosure of discovery materials would be willing to disclose absent the protective order; rather, the cases presume that since the only practical effect of the protective order is to prevent an otherwise willing speaker from communicating to a willing listener, the party seeking to intervene meets the redressable injury requirement of standing simply because the order impedes “the news agencies’ ability to discover newsworthy information from potential speakers.” *Davis v. East Baton Rouge Parish School Board*, 78 F. 3d 920, 927 (5th Cir. 1996).

The courts are in disagreement as to the burden of proof when motions are made to modify or vacate an existing protective order. One standard provides that, assuming the order to have been validly entered in the first instance, the moving party must show sufficient reasons to release the protected information. See *Phillips v. General Motors Corp.*, 307 F.3d 1206, 1213 (9th Cir. 2002). Another approach leaves the burden of proof on the party that sought the order in the first instance to justify continued confidentiality, but adds reliance on the existing order as a factor to be considered. See *Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 789-90 (3d Cir. 1994).

#### Best Practices

1. When intervention is allowed to oppose a motion for a protective order that has not yet been entered, or for purposes of challenging a general protective order, the court should consider and balance the public interest in the disclosure sought, the legitimate privacy interests that favor non-disclosure, and the extent to which the information is relevant to the controversy. The party seeking to limit disclosure has the burden of demonstrating that the balance of interests satisfies the “good cause” standard of Fed. R. Civ. P. 26(c).
2. If a protective order has already been entered after full consideration of the merits, including a review of the contents of the documents that are prohibited from disclosure under the order, the intervenor should be required to demonstrate circumstances or considerations not already considered by the court.
3. In entering into a confidentiality agreement or seeking a protective order, parties should anticipate non-party demands for discovery materials through subpoena, or that non-parties will object to producing information pursuant to subpoena because the requested information is claimed to be confidential. Agreements should also include provisions that expressly allow parties to provide materials to requesting regulatory agencies that offer appropriate protection for confidential materials.
4. When collateral litigants intervene, the issuing court “should satisfy itself that the protected discovery is sufficiently relevant to the collateral litigation that a substantial amount of duplicative discovery will be avoided...” *Foltz v. State Farm Mutual Automobile Ins. Co.*, 331 F.3d 1122, 1132 (9th Cir. 2003). The court should balance this policy of avoiding duplicative discovery against the countervailing interests of the parties opposing modification, including their reasonable reliance on the order’s nondisclosure provisions. In many cases, any legitimate interest in continued secrecy can be accommodated by placing the collateral litigants under the use and non-disclosure restrictions of the original protective order. Modification merely removes the impediment of the protective order in the collateral litigation. The collateral court retains authority to determine the ultimate discoverability of, and the protection to be afforded to, specific materials in the collateral proceedings. *Foltz*, 331 F. 3d at 1133.

## Examples

1. Using the example of the class action race discrimination lawsuit discussed under Chapter 1, Principle 1, a class member has opted out of the class and is pursuing an individual race discrimination lawsuit against the same defendant in another jurisdiction. The attorney for the opt-out plaintiff calls the attorney representing the plaintiff class and asks for a copy of the payroll and promotion database that defendant produced in the class action. However, this has been designated confidential, so class plaintiffs' counsel cannot provide the database to the attorney for the opt-out plaintiff.
2. In the same case as above, the attorney for the opt-out plaintiff serves a Fed. R. Civ. P. 45 subpoena on the attorney for the plaintiff class seeking production of all documents produced by defendant in the class action. But counsel for the plaintiff class is under a court order not to disseminate the confidential material produced by defendant. The protective order provides that class plaintiffs' counsel tenders the subpoena to defendant's counsel, who is obligated by the protective order to defend the protective order and oppose the subpoena, negotiate an extension of the protective order, or waive protection.
3. In the same case as above, because the opt-out plaintiff's lawsuit involves allegations similar to those involved in the class action, and direct discovery requests for the information would be inevitable, defendant's counsel agrees that the material can be produced to the opt-out plaintiff so long as he is willing to enter a similar protective order.
4. In a toxic tort case, a protective order is entered at the commencement of discovery. After significant discovery has taken place, a member of the news media moves to intervene, asserting that the general public has a legitimate interest in documents or information exchanged during discovery. If the court determines that a colorable public interest has been asserted, intervention should be allowed.
5. Assume the same facts as stated above in Example 4, except that the parties fail to agree upon a stipulated protective order and one party moves for entry of a protective order pursuant to Fed. R. Civ. P. 26(c). The media, asserting the public interest, should be permitted to intervene and be heard in opposition to the motion. The party asserting confidentiality bears the burden of proving that an order should issue under Fed. R. Civ. P. 26(c).
6. Assume the same facts as stated in Example 4, except that the parties to the litigation failed to agree on a stipulated protective order and one party moved for the entry of a protective order pursuant to Fed. R. Civ. P. 26(c). The judge ruled on the motion and issued an order, making appropriate findings of fact and conclusions of law. At a later date, the news media intervene, seeking access to documents subject to the order. The intervenors bear the burden of coming forward with evidence sufficient to overcome the initial presumption that the existing order remain in place, although the original proponent bears the ultimate burden of persuasion that the order continues to be necessary and narrowly tailored to protect legitimate privacy and confidentiality interests.



## CHAPTER I Selected Bibliography

*Manual for Complex Litigation, Fourth* (2004), Sec. 11.432 (discussing the use of “umbrella” and “particularized” protective orders in complex litigation).

*Florida Statutes Annotated*, §69.081 (describing the procedure for entry of an order “which has the purpose or effect of concealing a public hazard,” etc.).

*Texas Rule of Civil Procedure 76a* (declaring that certain court records are presumed to be open and setting forth standards for sealing same).

*Seattle Times v. Rhinehart*, 467 U.S. 20 (1984) (no public right of access to unfiled discovery materials).

*Poliquin v. Garden Way*, 989 F.2d 527 (1st Cir. 1993) (showing of “good cause” sufficient to protect discovery materials from disclosure insufficient to bar access to evidence introduced at trial).

*SEC v. TheStreet.com*, 273 F.3d 222 (2d Cir. 2001) (transcripts of depositions of NYSE officers held not to be “judicial documents” but no presumption against public access as third parties had been present).

*Leucadia v. Applied Extrusion Tech., Inc.*, 998 F.2d 157 (3d Cir. 1993) (recognizes distinction for purposes of access between discovery materials filed as exhibits to “merits” motions and those filed in support of discovery motions).

*Shingara v. Skiles*, 420 F.3d 301 (3d Cir. 2005) (reviewing standards for issuance of umbrella orders and concluding such an order was not warranted in a case that “is neither complex nor involves ‘large-scale discovery’” and good cause showing not made).

*Pittston Co. v. United States*, 368 F.3d 385 (4th Cir. 2004) (holds that party cannot object to sealing of materials when it had access to same and had agreed to sealing).

*Procter & Gamble Co. v. Bankers Trust Co.*, 78 F.3d 219 (6th Cir. 1996) (protective order cannot allow parties to control access to judicial records).

*Baxter Internat'l v. Abbott Labs*, 297 F.3d 544 (7th Cir. 2002) (distinguishes between broad procedural orders entered for purposes of discovery as opposed to those entered at later stages of litigation).

*Citizens First Nat'l Bank v. Cincinnati Ins. Co.*, 178 F.3d 943 (7th Cir. 1999) (broad protective order granting parties “a virtual carte blanche” to seal records invalid under Fed. R. Civ. P. 26).

*In re Remington Arms Co.*, 952 F.2d 1029 (8th Cir. 1991) (sets forth procedure to determine whether information sought in discovery is trade secret for purposes of Fed. R. Civ. P. 26(c)(7) protective order).

*United Nuclear Corp. v. Cranford Ins. Co.*, 905 F.2d 1424 (10th Cir. 1990) (discusses access by collateral litigants to discovery materials subject to protective order).

*Foltz v. State Farm Mut. Auto Ins. Co.*, 331 F.3d 1122 (9th Cir. 2003) (discusses access by collateral litigants to discovery materials subject to protective order; establishes “mechanics of the relevance inquiry” and which court should do what).

*United States v. Microsoft Corp.*, 165 F.3d 952 (D.C. Cir. 1999) (depositions in an antitrust action are open to public pursuant to the Publicity in Taking Evidence Act).

*A full bibliography, updated periodically, may be found on The Sedona Conference web site at [www.thesedonaconference.org](http://www.thesedonaconference.org)*

## CHAPTER 2

### PLEADINGS, COURT ORDERS, SUBSTANTIVE MOTIONS AND DOCKETS

**Principle 1** In civil proceedings, the public has a qualified right of access to documents filed with a court that are relevant to adjudicating the merits of a controversy. In compelling circumstances, a court may exercise its discretion to deny public access to submitted documents to protect the privacy, confidentiality or other rights of the litigants.

Chapter 1 dealt with discovery materials exchanged between the parties, for which there is no qualified right or presumption of public access. In Chapter 2 we start with the act of filing a document with the court, such as a pleading, response, motion, or an exhibit. We also deal with the documents and records created by the court in relation to civil litigation, such as dockets, docket entries, memoranda, orders, or judgments. When a document is filed or created by the court, a qualified right or presumption of public access arises. However, this is just the starting point of the public access analysis. The strength of the presumption, and the consequent burden that must be met to overcome it, depends on the relationship of the document to the adjudicative process. The more important the document is to the core judicial function of determining the facts and the law applicable to the case, the stronger the presumption and the higher the burden to overcome it. Thus, a qualified right or presumption of public access attaches to all documents filed with the court that are material to the adjudication of non-discovery matters. See *In re Cendant Corporation*, 260 F.3d 183, 192-93 (3d Cir. 2001); *Chicago Tribune Co. v. Bridgestone/Firestone, Inc.*, 263 F.3d 1304, 1312 (11th Cir. 2001).

Courts have found the qualified right or presumption of public access from either of two sources, the Constitution or the common law. The United States Supreme Court has found a First Amendment right of public access in criminal cases. A constitutional right to public access arises if the proceedings or documents have historically been open to the general public and “public access plays a significant positive role in the functioning of the particular process in question.” *Press-Enter. Co. v. Superior Court*, 478 U.S. 1, 8 (1986). “If the particular proceeding in question passes these tests of experience and logic, a qualified First Amendment right of public access attaches.” *Id.* That qualified right can be overcome “only by an overriding interest based on findings that closure is essential to preserve higher values and is narrowly tailored to serve that interest.” *Id.* See also *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 606-607 (1982) (“Where, as in the present case, the State attempts to deny the right of access in order to inhibit the disclosure of sensitive information, it must be shown that the denial is necessitated by a compelling governmental interest, and is narrowly tailored to serve that interest.”).

Some federal courts have extended the First Amendment right of public access to court files and proceedings in civil cases. See, e.g., *Republic of Philippines v. Westinghouse, Elec. Corp.*, 949 F.2d 653, 659 (3d Cir. 1991); *Rushford v. New Yorker Magazine, Inc.*, 846 F.2d 249, 252-53 (4th Cir. 1988). Other courts have found a common law presumption in favor of public access for documents and proceedings in civil cases. See, e.g., *FTC v. Standard Fin. Mgr.*, 830 F.2d 404, 408 n.4 (1st Cir. 1987); *Brown & Williamson Tobacco Corp. v. FTC*, 710 F.2d 1165, 1179 (6th Cir. 1983); *Smith v. United States District Court*, 956 F.2d 647, 650 (7th Cir. 1992); *EEOC v. Erection Co.*, 900 F.2d 168, 169 (9th Cir. 1990); *NBC Subsidiary (KNBC-TV), Inc. v. Superior Court*, 20 Cal. 4th 1178, 1217 (1999). Some courts have found both a First Amendment right and a common law presumption. See e.g., *Lugosch v. Pyramid Co. of Onandaga*, 435 F.3d 110 (2d Cir. 2006) (Documents submitted to a court for its consideration in a summary judgment motion are “judicial documents” to which a presumption of immediate access applies under both the common law and the First Amendment.)

Public access serves important public interests, including informing the public of the cases that are before the court. The essential benefit of access is to ensure accountability to the public. Public scrutiny brings about accountability, fostering public trust that the judicial system is functioning justly, properly and efficiently, and that participants in the system are properly and honestly performing their duties. Another benefit of public access is educating the public on the

workings of the courts and the civil justice system, which promotes public confidence in the judicial system.

While important interests are served by the First Amendment right or common law presumption of public access, the right is not absolute and the presumption may be overcome in appropriate circumstances. If the documents in question have little or no relation to the merits of the case or have not historically been available to the public, the presumption of public access that arises from the mere fact that these have been filed with the court is quite weak. For instance, if the documents have been filed in relation to a discovery dispute, the court may seal them under the same “good cause” standard that would support a particularized protective order under Fed. R. Civ. P. 26(c). Likewise, the presumption of public access to filed settlement documents that merely recite the terms of settlement and do not purport to assign liability or otherwise “adjudicate” the case is weak. See *Gambale v. Deutsche Bank AG*, 377 F.3d 133, 143-44 (2d Cir. 2004) (distinguishing between settlement amount and summary judgment documents).

The presumption of public access is much stronger in relation to “judicial documents,” that is, documents that relate to the merits of the case and assist the court in fulfilling its adjudicatory function. Courts have the discretion to seal court documents to protect rights of privacy or confidentiality, provided the court determines that an “overriding interest” exists that can only be protected by such an order. See, e.g., *Chicago Tribune Co. v. Bridgestone/Firestone*, 263 F.3d 1304, 1313 (11th Cir. 2001); *In re Cendant Corp.*, 260 F.3d at 194; *Leucadia, Inc. v. Applied Extrusion Tech., Inc.*, 998 F.2d 157, 166 (3rd Cir. 1993); *NBC Subsidiary (KNBC-TV), Inc. v. Superior Court*, 20 Cal. 4th 1178, 1217 (1999). A circumstance that might justify the exercise of the court’s discretion is to protect the confidential proprietary interests in trade secrets or other commercially sensitive information. See *Leucadia*, 998 F.2d at 166; *Joint Stock Soc’y v. UDV North America*, 104 F. Supp. 2d 390, 396 (D. Del. 2000). Whether or not the requisite need for sealing has been demonstrated is a matter of courts’ supervisory powers and is “best left to the sound discretion of the court, discretion to be exercised in light of the relevant facts and circumstances of the particular case.” *Nixon v. Warner Communications*, 435 U.S. 589, 598-99 (1978).

During the 18 month period between the publication of the “public comment” draft of this document in April 2005 and the final draft at the end of 2006, several commentators put forward the proposition that all actions to restrict or deny public access to the documents and proceedings related to civil litigation should be subject to the same “good cause” showing to overcome the presumption in favor of public access. While this notion has simplicity and surface appeal, it is contradicted by the great weight of the case law and advanced in a limited circumstance by only one federal circuit court decision. *Chicago Tribune Co. v. Bridgestone/Firestone*, 263 F. 3d 1304, 1313 (11th Cir. 2001) (holding that the “good cause” standard of Fed. R. Civ. P. 26(c) may also apply to restrict access to documents filed in relation to a failed summary judgment motion in a settled case). While all justifications for restricting public access must constitute “good cause” to be upheld, the requisite “good cause” will be dramatically different depending upon the particular documents and proceedings, and the particular stages of the litigation. The “good cause” needed to support an umbrella protective order under Fed. R. Civ. P. 26(c) will not suffice to support the sealing of those discovery documents when filed in court. To lump the varying expressions of burden under the common rubric of “good cause” creates confusion as to what standard actually applies, and does a disservice to attorneys and parties who may erroneously believe that the “good cause” that supported an umbrella protective order in discovery will also support an order to seal a record on appeal. See e.g. *Baxter International v. Abbott Labs*, 297 F. 3d 544 (7th Cir. 2002).

Statutes and court rules recognize specific situations in which pleadings must be kept confidential. For example, under the *qui tam* provisions of the False Claims Act, 31 U.S.C. §3729 et seq., a complaint of a private person must be filed *in camera* and remain sealed for at least 60 days, and the government may move *in camera* for an extension of the sealing. 31 U.S.C. §§3730(b)(2) and (3). The purpose of the sealing requirement is to give the government an opportunity to review the complaint and determine whether to intervene. 31 U.S.C. §3130(b)(4) Similarly, the Trademark Counterfeiting Act of 1984, codified in relevant part at 15 U.S.C. Sec. 1116(d), provides that a court may, on *ex parte* application, issue a seizure order for goods and

counterfeit marks, and that any such order, “together with the supporting documents,” shall be filed under seal, “until the person against whom the order is directed has an opportunity to contest such order...” 15 U.S.C. §1116(d)(8). The purpose of the sealing is to avoid the loss or concealment of the items to be seized. 15 U.S.C. §1116(d)(4)(B)(vii). In both of these statutory examples, sealing is for a specific purpose and a limited time.

There are also categories of cases with records that invariably include information, the disclosure of which is discouraged as a matter of public policy. State legislatures have adopted statutes mandating that such categories of cases be closed to the public. *See, e.g.*, Indiana Code §31-39-1-2 (records in juvenile proceedings); Maryland Code, State Gov’t. Art., §10-616(b) (records in adoption and guardianship proceedings); Maryland Code, Art., §10-616(b) (records in adoption and guardianship proceedings); Maryland Code, Art. 88A, §6(b) (records in child abuse or neglect proceedings); New York, Family Court Act, §166 (records of family court proceedings); New York, Domestic Relations Law, §235 (records in divorce, custody, and child support proceedings). Several state courts have, or are considering, court rules restricting remote access via the Internet in certain categories of cases. *See, e.g.*, Indiana Administrative Rule 9(g); Maryland Rules 16-1006 and 16-1007; Vermont Rules for Public Access to Court Records, §6. The Judicial Conference of the United States has also determined that remote access via the Internet to at least one category of cases, Social Security appeals, should be restricted. *See* Report of the Judicial Conference Committee on Court Administration and Case Management on Privacy and Public Access to Electronic Case Files, September 2001.

A court might also determine that sealing an entire document is not necessary because the protection of privacy or other interests can be achieved by redacting the private information or allowing the use of fictitious names. In such cases, redaction or use of a pseudonym can preserve the important privacy rights or property interests, while protecting the public’s right to access. *See Doe v. City of Chicago*, 360 F. 3d 667, 669-670 (7th Cir. 2004). In other cases, sealing may be warranted only for a limited period of time. *See* E-Government Act of 2002, Public Law 107-347, and proposed Fed. R. Civ. P. 5.2, (implementing the E-Government Act, §205(c)(3), to provide privacy protection for filings electronically transmitted to a court), approved by the Civil Rules Advisory Committee of the U.S. Judicial Conference. *See* [http://www.uscourts.gov/rules/Rules\\_Publication\\_August\\_2005.pdf#page=55](http://www.uscourts.gov/rules/Rules_Publication_August_2005.pdf#page=55)

Courts should adopt procedures to facilitate the orderly consideration of motions related to public access to filed documents. For example, a party may believe it has compelling reasons to request an order allowing it to file an entire document with the court under seal. The court may permit the party to “lodge” the document with the court (a step short of filing the document) and file an appropriate motion to have the document sealed upon filing. Under such a procedure, the court and opposing parties have access to the document while the motion is under consideration, and the public has notice of the pending motion to file the document under seal, but the document has not been made public through filing. *See, e.g.*, D.N.J. L. Civ. R. 5.3; Cal. R. Ct. 243.1 to 243.4

#### Best Practices

1. Courts should provide guidance to civil litigants regarding the procedures for seeking and the standards for obtaining protective orders and sealing orders through the promulgation of local rules, court-wide standing orders, and courtroom-specific standing orders.
2. In jurisdictions where lodging is permitted, litigants should temporarily file or “lodge” documents with the court awaiting a ruling on a motion to have the documents filed under seal. If the motion is granted, the documents will be filed under seal. If the motion is denied the documents will be returned to the party unfiled.
3. Whenever a party seeks to file documents previously subject to a protective order, appropriate notice should be given that would enable the producing party to move for protection of the documents.

4. In exercising its discretion to issue a sealing order, a court should consider the importance of the rights and interests (for instance, privacy rights, confidential information, or proprietary business information) that would be jeopardized by public access to the sealed material; whether the need for confidentiality outweighs the public's interest in disclosure; whether the sealing requested is broader than necessary to meet that need for sealing; and whether less restrictive alternatives might be available that would preserve the interest in confidentiality while permitting at least some public access such as redaction, limiting the duration of the sealing order or the use of pseudonyms. Absent extraordinary circumstances, sealing orders themselves should not be sealed.
5. Attorneys should counsel their clients regarding the standards for sealing party-filed documents, the risk that information may be made public, and measures that may be taken to minimize or avoid the compromise of such information.
6. Sealing requests should not be overly broad and attorneys should take reasonable steps to segregate material that should be filed under seal from material that may be filed without seal. For example, an entire document a party requests to file under seal should not be sealed if, as a practical matter, confidentiality can be adequately protected by more limited means, such as the redaction of specific information, the use of "Doe" pleadings, or the sealing of only a portion of the document, with the non-confidential material to be openly filed.
7. Orders granting or denying a motion to seal should be subject to immediate appellate review under the collateral order or other doctrine.

#### Examples

1. A plaintiff files a trademark infringement action alleging that the defendant is manufacturing "knock off" products. The plaintiff seeks an *ex parte* seizure order, pursuant to statute, based on a showing that the defendant intends to move or destroy the products. The complaint and the request for an immediate seizure order are temporarily filed along with a motion to have the documents filed under seal. The temporary filing is accompanied by a proposed order to be endorsed by the judge, also under seal, stating that the complaint will be unsealed upon either the execution or the denial of the requested seizure order.
2. The exclusive Midwestern distributor of a highly popular and profitable product of an East Coast manufacturer is engaged in a dispute with the manufacturer over the manufacturer's alleged shoddy accounting practices. For several months the parties engaged in unsuccessful settlement negotiations, until the distributor prepares to file suit. The plaintiff distributor requests that its attorney file the complaint under seal to avoid publicity and loss of good will. After review of the controlling law in the jurisdiction, the attorney determines that the plaintiff's generalized interest in avoiding publicity and losing goodwill will not provide the requisite showing to support a sealing order. The attorney so informs his client and counsels him regarding possible alternatives, including alternative dispute resolution.
3. The same facts as Example 2 above. The distributor files a motion for summary judgment supported by documents designated as "confidential" and produced under a Rule 26(c) protective order which detail proprietary and commercially sensitive information about the manufacturer's business. The distributor's attorney files the motion with the confidential information redacted and requests the court's permission to file an unredacted version under seal. The manufacturer's attorney promptly files a motion explaining the basis for sealing the unredacted version with an affidavit that establishes the significance of the information to the business, the competitive harm disclosure might cause, and the measures taken to date to keep the information confidential. The court grants the sealing motion and files an order, stating the grounds for sealing without revealing any of the information itself.
4. Several members of a wealthy family, all beneficiaries under a trust, file suit against another family member who is the trustee, alleging misappropriation of trust assets. In the complaint, the plaintiffs set forth sensitive family information, such as the value of the trust, their respective

shares of the trust assets, and the holdings of the trust. The plaintiffs' attorney prepares two versions of the complaint: one version in which the paragraphs containing the sensitive family information have been redacted, to which the public would have access, and another version of the complaint to be the subject of a motion to seal.

5. A plaintiff who is HIV-positive brings suit against a dentist for refusal to treat him in violation of state law. The plaintiff is not known in his community as being HIV-positive. He files his complaint under the court's "Doe" procedure, in which information that would identify him personally is removed from the version of the pleading to which the public has access. The version of the complaint containing identifying information is filed under seal. The judge issues an order that information identifying the plaintiff in all subsequent pleadings is to be redacted from publicly accessible versions, with unredacted versions filed under seal.
6. A plaintiff intends to file a motion for summary judgment in fourteen days, which will be supported by documents it obtained from the defendant through discovery and designated by the defendant as confidential under a stipulated protective order. Pursuant to the terms of the particular protective order, the plaintiff informs the defendant of its intention to file, which permits the defendant to file a motion to seal the material in advance of the summary judgment filing date. The defendant files the motion to seal, and the plaintiff files a redacted version of its summary judgment motion and supporting information but separately files an unredacted version containing the confidential information with the court.
7. During the course of litigation, the court indicates it may impose sanctions on counsel for repeatedly missing court-imposed deadlines in the case. Counsel requests the court's permission to file a responsive brief and affidavit under seal simply explaining that "medical reasons" are detailed in its response regarding sanctions. The court exercises its discretion to seal the brief and affidavit, in part because the public's interest in access is weaker because the filings have no role in the determination of the merits of the case.
8. In the same matter, counsel files a dispositive summary judgment motion, and temporarily files select portions of the supporting evidence along with a motion to seal stating that "confidential and proprietary" business information is contained in the documents. The court indicates its initial intent to deny the sealing motion. Counsel then requests, and the court grants, permission for counsel to file a supplemental brief and affidavit establishing how the information in the documents is maintained in confidence, that it is not known to the public or business competitors, and how disclosure of the information may give competitors an unfair advantage. In light of the supplemental filings, the court assesses the strength of the presumption of public access, weighs it against the interests in confidentiality, and finds a compelling need to partially seal the documents.

**Principle 2    The public has a qualified right of access to court dockets that can only be overcome in compelling circumstances.**

The docket is the principal index of judicial proceedings. All the judicial business of the court should be noted on the docket. See Fed. R. Civ. P. 79(a). For each case, the individual docket should serve as a record of all activity and as an index of all documents, pleadings, appearances, the scheduling of hearings and trials, motions, orders, judgments, as well as miscellaneous items.

There is a presumption of public access to dockets. See, e.g., *Hartford Courant Co. v. Pellegrino*, 380 F.3d 83, 93-94 (2nd Cir. 2004). Access to the docket is the primary means to determine if a particular case is being adjudicated or if a party is engaged in litigation. Access to individual case dockets is the primary means to monitor the course of any particular case. Access to the docket is also the primary means for the public (including the media, academics, and civic groups) to monitor the overall performance of the courts and the administration of justice. The effectiveness of particular laws or court rules is often measured by analysis of court dockets. Moreover, legislative decisions regarding the allocation of resources to the judicial branch of government are based in large part on statistical analysis of docket activity.

When permitted by law or statute, certain court documents may be sealed and certain proceedings may be closed to the public. However, it does not follow that corresponding dockets should be sealed, either in whole or in part. The existence of a case itself should never be kept secret, and whenever particular documents or proceedings are to be sealed, docket entries referencing that sealing should be made to give the public adequate notice.

On rare occasions, docket entries could reveal information that would jeopardize the privacy or confidentiality interests of parties involved. Statutes and court rules restrict public disclosure of particular types of information that might appear in docket entries, such as Social Security numbers or the names of minors. However, the restrictions imposed by such statutes and court rules do not justify the sealing of the docket itself.

Rare circumstances may justify the temporary redaction of particular information in docket entries to prevent the destruction of evidence or the loss of a remedy. When such redaction is required a judge should make appropriate findings of fact and conclusions of law in an order that should be noted on the docket and should expire by its own terms when the circumstances justifying the order have passed.

Many courts include narrative “minute entries” or summaries of proceedings directly on the docket when the proceeding generates no document. Clerks who compose such narrative entries should be careful not to include sensitive or confidential information explicitly restricted by statute, court rule, or order. A narrative entry that must be redacted or sealed should be composed as a separate document and placed in the case file, and an appropriate entry made on the docket.

#### Best Practices

1. The public should be given access to the docket except in the most compelling circumstances. Even if the merits of a case warrant issuance of a sealing order, the case must still be assigned a docket number and a docket index itself must be accessible to the public.
2. The identity of the judge to whom a case has been assigned should appear on the docket under all circumstances, as should the identity of counsel.
3. “Procedural” events should appear on the docket for public review. These include, for example, the nature of the case, the payment of the filing fee, and notations that motions have been made or affidavits filed.
4. There may be circumstances under which parts of individual docket entries should be sealed. If so, the existence of a sealing order based on findings of fact and conclusions of law should be reflected on the docket.
5. Care should be exercised to distinguish between situations in which a document or proceeding is sealed, and the extremely rare instances in which the corresponding docket entries should also be sealed.

#### Examples

1. A plaintiff seeks to file a trademark infringement action against a defendant and seeks an immediate *ex parte* seizure order. The plaintiff is aware that the defendant monitors the filing of suits against it. The plaintiff presents the complaint and the *ex parte* motion for a seizure order, arguing that if the defendant knows the existence of the litigation it will destroy or transfer the infringing products to an unknown location. Convinced that the standards for both the *ex parte* seizure order and a sealing order are met, the judge endorses the seizure order and dictates a temporary sealing order, reciting her findings of fact and conclusions of law. The complaint and orders are filed with the clerk, with instructions to assign the case a number and docket it with only the number, judge’s name, entries for a complaint and seizure order under seal, and an entry for a temporary sealing order. The temporary sealing order is set to expire upon execution of the seizure order or after a reasonable period of time.

2. An infant plaintiff, by his guardian, commences a declaratory judgment action against an insurer that issued a homeowners' policy to the infant plaintiff's grandfather. The infant alleges that, while visiting his grandfather's home on various occasions, he was sexually assaulted by the grandfather. The infant seeks a declaration that the grandfather's acts fall within the policy coverage. The case is assigned a "John Doe" plaintiff name, in accordance with local practice regarding the identity of minors. The insurance company is named as defendant. The docket includes the names of counsel and of the judge and routine procedural entries. The complaint appears on the docket with the notation "Under Seal," followed by an entry for the sealing order. The sealing order, which is not itself under seal, states that the judge, after reviewing the complaint and hearing from counsel, finds that the allegations in the complaint involve a minor and include matters of a sensitive and private nature, which should be confidential and protected from public disclosure by the court and counsel, in the best interests of the child.
3. A plaintiff in a patent infringement action has moved for summary judgment against the defendant. As part of that motion, the plaintiff has submitted the affidavit of an expert economist containing financial information that is proprietary to the plaintiff and which, if made public, would have adverse competitive effects on the plaintiff. The motion for summary judgment is filed and the affidavit is filed separately under seal. Both filings are noted in the docket, including the name of the affiant expert.

**Principle 3 There is a qualified right of access to judgments, judicial opinions and memoranda, and orders issued by a court that can only be overcome in compelling circumstances.**

The public has a qualified right of access to judgments, judicial opinions and memoranda, and court orders. Public access to opinions and orders is essential to public understanding and monitoring of the judicial process. See *Republic of the Philippines v. Westinghouse Electric Corp.*, 949 F.2d 653, 663-64 (3d Cir. 1991); *In re Continental Illinois Securities Litigation*, 732 F.2d 1302, 1308-09 (7th Cir. 1984). Access by the legal community to court opinions is essential to the development of the common law and *stare decisis*, which depends on the application of precedent as expressed in court orders, judgments, and the reasoning articulated in judicial opinions and memoranda. Absent a clear showing of a compelling interest that can only be protected by imposing confidentiality restrictions, access to judgments, judicial opinions and memoranda and court orders should not be restricted. If restrictions on access are found to be essential, those restrictions should be narrowly crafted so as to impose no greater restriction on public access than is necessary to protect the interest at stake.

Courts and attorneys must balance the need for access with the legitimate interests of the parties in privacy and confidentiality. Court documents should be written with the presumption of public access in mind, and with the understanding that *de facto* restrictions on access arising from the difficulties of retrieving physical files have been largely eliminated by technology. Consequently, judges and law clerks should not include unnecessary confidential information in filed documents. Attorneys who submit proposed orders for a judge's signature should also exercise similar editorial judgment.

When information restricted by a statute or court rule must be included in a document, or when courts are presented with confidential information that must be incorporated in rulings, an unredacted version should be filed under seal and made available to the attorneys and parties in the case. A redacted version should be filed for public access, with an explanation of the reason for the redactions. The sealing of an entire document, or the filing of a redacted document under circumstances not addressed by a statute or court rule, must be based on a finding by the judge, available in the public record, that the necessity for restricting public access to the document in question overcomes the presumption of public access to judicial decisions in that particular instance.



### Best Practices

1. Judges should include in opinions and orders a full discussion of the facts relevant to their decisions, guided by the consideration that the purpose of opinions and orders is to provide to litigants, lawyers, the public, and appellate courts reasoned explanations for their decisions. Courts should avoid including information, the public disclosure of which may be harmful to a litigant or a third party. If there is a compelling reason to include such information in opinions or orders, courts should consider redacting or sealing the information the disclosure of which would be harmful to a litigant or third party.
2. When a court includes information restricted by statute, court rule, or an existing sealing order in an opinion or order, the court should issue both a redacted document, as to which there will be public access, and an unredacted document that will be filed under seal.
3. If an opinion, memorandum, order, or judgment is to be sealed in part or in whole, the judge should set forth findings of fact and conclusions of law that will be available for public access, including a discussion of the right deemed to be threatened by public access to the full opinion, the lesser restrictive alternatives to sealing the court considered, and why these were rejected.
4. The court should review its decision to seal an opinion, memorandum, or order, upon request by the parties or a third party intervenor, or upon the occurrence of changed circumstances. The sealing order should be vacated if and when the grounds justifying the sealing order no longer exist.

### Examples

1. A business entity commences an action against a competitor for the theft of information utilized by the plaintiff in a manufacturing process. The plaintiff's manufacturing process is proprietary and is a closely guarded secret. For the purpose of ruling on a motion for summary judgment, the judge must compare both parties' manufacturing processes in an opinion. The judge files a brief summary judgment order for public access, together with a finding that the manufacturing process constitutes a legally protected trade secret and that public access to the accompanying opinion would compromise the plaintiff's legal interest. The opinion is filed under seal.
2. Same facts as Example 1 above. The theft is alleged to have been committed by a former employee of the plaintiff who became employed by the defendant, and the defendant has not yet established its own manufacturing process. The relief sought by plaintiff is solely injunctive in nature. The judge, ruling on a motion for a preliminary injunction, addresses whether the former employee had access to the plaintiff's information and the likelihood of the plaintiff's success on the merits of the case, but carefully avoids describing the information in detail in the opinion and order.
3. Same facts as Example 1 above. The opinion issued by the judge totals thirty pages addressing a number of facts and legal issues, and the detailed comparison of the parties' manufacturing processes appears on only three pages. A sealing order covering all information about the manufacturing process, based on findings of fact and conclusions of law, is already in place in the case. The judge files the full opinion under seal and a redacted opinion, referencing the existing protective order, for public access.
4. An insurance broker brings an action against an insurer, alleging that she is owed commissions on the sale of the insurer's products, and that other brokers were paid at a higher rate than she was. During discovery, the parties stipulate to a protective order for the production of records of commissions paid by the defendant to other brokers. Both parties move for summary judgment and include in their papers the records of commissions paid to other brokers. In ruling on the motions, the judge does not treat the commission information as confidential, as the protective order applied only to non-filed discovery materials and no showing has been made to justify sealing.

**Principle 4 Notice of motions to seal and supporting materials should be reflected in the publicly accessible docket.**

The public has a qualified right of access to all papers, including briefs, submitted in support of and in opposition to motions filed with the court that address non-discovery matters. Consequently, motions to seal documents should be rare and made only when there is no feasible alternative. However, in some circumstances, papers filed with the court may contain information that implicates legitimate privacy or confidentiality interests, or otherwise warrants redaction or sealing, and may be proper subjects for a sealing order before filing. In the absence of court rules or orders specifying procedures for the filing of materials under seal, the parties themselves should agree to a procedure under which a party intending to publicly file papers or information which has been subject to a sealing or protective order, or intending to file under seal any new material, informs the other party in a timely manner of that intent, thus allowing the other party opportunity to object, consent, or confer on the appropriate confidentiality status of the material. A court is not required to give particularized notice to any specific constituency when deciding a motion to seal. *See Webster Groves Sch. Dist. v. Pulitzer Publ'g Co.*, 898 F.2d 1371 (8th Cir. 1990).

**Best Practices**

1. As stated under Principle 1, attorneys should take reasonable steps to minimize the use of confidential information in motions, proposed orders, or briefs, so as to minimize the need to redact or seal such documents. The inclusion of attachments, such as discovery material produced under a protective order, should likewise be minimized, and when necessary, should be presented in such a manner as to facilitate the temporary filing of only the protected attachments, and not the whole document, pending resolution of a motion to seal.
2. Attorneys who intend to submit a motion discussing or attaching material designated by another as confidential or subject to seal should inform the designating party of their intent to file the material and ask the designating party to promptly respond and confirm whether that party continues to contend that the material merits confidentiality.
3. For cases in which the parties are unable to agree upon what, if any, information should be included in any public filing, and in all cases in which new pleadings, substantive motions, and accompanying materials are submitted with the intention of being filed under seal, the court should permit (by way of local rule, standing order, or in a protective order under Rule 26(c)) a party to temporarily file the material with the court under seal, thereby providing notice to all parties of the intended action without compromising the purported private or confidential nature of the material itself, pending a prompt ruling on confidentiality.

**Examples**

1. A plaintiff intends to file a motion for summary judgment in fourteen days, which will include documents obtained from the defendant through discovery and designated by the defendant as confidential under a stipulated protective order. The plaintiff may inform the defendant of its intention to include specified documents in an upcoming filing and request that the defendant file a motion to seal the material in advance of the filing date. If the defendant files the motion to seal, the plaintiff temporarily files its summary judgment motion and the accompanying documents with the court, pending a decision on the defendant's motion to seal. If the defendant does not move to seal in advance of the filing, the plaintiff still temporarily files its summary judgment motion and the accompanying documents but also files a motion seeking a court order that the summary judgment motion be filed publicly and that defendant's confidentiality designation be removed. If the defendant does not respond or object, defendant will be deemed to have waived the confidential designation, the designation will be removed, and the summary judgment motion with the supporting documents will be publicly filed.
2. The same facts as above, except that the plaintiff agrees that the documents merit sealing. The plaintiff drafts its summary judgment motion, brief, and supporting declarations, limiting its use

of confidential material and reference to confidential facts. It temporarily files its motion and accompanying materials with the court along with a motion to seal, and submits a redacted copy of the summary judgment motion and brief intended for public filing. The court reviews the documents proposed to be sealed and the documents proposed to be redacted, and rules upon the motion to seal.

**Principle 5 Non-parties may seek leave to intervene in a pending case to oppose a motion to seal, to have an existing sealing order modified or vacated, or to obtain a sealing order.**

Courts have recognized that the members of the public have standing, and grounds to intervene, to obtain access to documents filed with a court under seal. So too, non-parties who have an interest in privacy and confidentiality of materials filed with the court and available to the public also have standing to independently seek to seal such materials. When the parties agree to secrecy or limitations on disclosure based upon interests that may be narrower than those of non-party intervenors, the court is not likely to have the benefit of the adversary process in making its decision. In such circumstances, courts have the discretion to grant non-parties permission to intervene for the purpose of opposing a pending motion to seal, or moving to have an existing sealing order modified or vacated. *Cf. Chiles v. Thornburgh*, 865 F.2d 1197, 1213 (11th Cir. 1989). The proposed intervenor should demonstrate good cause for the intervention and that it has moved to intervene without undue delay.

The court issuing a sealing order retains jurisdiction to entertain motions to modify or vacate the order after a case is closed. *See Gambale v. Deutsche Bank AG*, 377 F.3d 133 (2d Cir. 2004) (court retains jurisdiction over its own records, and may modify or vacate a sealing order after a case is closed.).

Appellate review exists for a trial court order addressing a motion to intervene and to modify or vacate a sealing order, although federal courts are split as to whether the proper “vehicle” is a *mandamus* petition or an appeal under 28 U.S.C. §1291. *See Virginia Dept. of State Police v. Washington Post*, 386 F.3d 567, 574 n.4 (4th Cir. 2004); *United States v. McVeigh*, 119 F.3d 806, 809-810 (10th Cir. 1997); *Bank of America Nat’l Trust and Savings Ass’n v. Hotel Rittenhouse Assoc.*, 800 F.2d 339, 341 n.2 (3d Cir. 1986). Whatever the vehicle, “all circuits that have considered the issue have found appellate jurisdiction ... under one doctrine or the other.” *McVeigh*, 119 F.3d at 810.

#### Best Practices

1. The entry of a motion to seal on the docket reasonably in advance of a hearing or decision provides adequate public notice that the court may seal party filed documents,
2. Interested persons who make an appropriate showing under Fed. R. Civ. P. 24(b) should be permitted to intervene to oppose, seek modification of, or seek vacatur of a sealing order.
3. Orders regarding motions to intervene and motions to modify or vacate sealing orders should be subject to immediate appellate review under the collateral order doctrine or any other appropriate procedural mechanism.

#### Examples

1. Company A wishes to file a motion for summary judgment including financial materials it obtained through discovery pursuant to a protective order. Company A files a motion to seal in advance of or concurrently with its motion for summary judgment. The allegedly confidential material is temporarily filed under seal with the court pending a decision on the motion to seal. The motion to seal and the temporary filing under seal of the financial materials appear on the court docket. The local newspaper, which has been reporting on the case, moves to intervene to oppose Company A’s motion to seal. The motion to intervene should be granted.

2. Same facts as in Example 1, and Company A obtains the order sealing material in connection with its summary judgment motion. Several months later a litigant in another action involving Company A discovers the sealing order and believes that the subject financial materials are relevant to the new action. The litigant's attorney moves to intervene and unseal the records. The court grants the motion to intervene. Subsequently, the court may, in its discretion, grant or deny the collateral litigant access to the sealed materials.

## CHAPTER 2

### Selected Bibliography

*Hartford Courant Co. v. Pellegrino*, 380 F.3d 83 (2d Cir. 2004) (addresses common law right of access to docket sheets).

In re *Application of Dow Jones & Co., Inc.*, 842 F.2d 603 (2d Cir.), cert. denied, sub nom. 488 U.S. 948 (1988) (entry of a protective order may give rise to redressable injury).

*North Jersey Media Group, Inc. v. Ashcroft*, 308 F.3d 198 (3d Cir. 2002) (analyzes media claim of access to “special interest” deportation proceedings under “experience and logic” test of *Richmond Newspapers*, 448 U.S. 555 (1980), and denies access; notes that some courts of appeals have extended First Amendment right to civil trials).

*Virginia Dept. of State Police v. Washington Post*, 386 F.3d 567 (4th Cir. 2004) (discusses distinction between First Amendment and common law rights of access and between summary judgment and discovery materials for purposes of access; discusses criminal investigation as compelling interest).

In re *Washington Post Co.*, 807 F.2d 383 (4th Cir. 1986) (extends common law right of access to documents submitted with regard to proceedings that fall within that right; reminds district courts to follow “procedural requirements” set out in earlier circuit decision).

*Detroit Free Press v. Ashcroft*, 303 F.3d 681 (6th Cir. 2002) (conducts same analysis as *North Jersey Media Group*, cited above; reaches opposite conclusion).

*Webster Grove School Dist. v. Pulitzer Publishing Co.*, 898 F.2d 1371 (8th Cir. 1990) (upholds limitation on access to civil proceeding involving juveniles; does not decide whether First Amendment right of access extends to civil proceedings, but applies common law standards; holds that procedural information on docket should not have been sealed).

*Chicago Tribune Co. v. Bridgestone/Firestone, Inc.*, 263 F.3d 1304 (11th Cir. 2001) (discusses right of access to sealed materials under First Amendment, common law and Rule 26(c); distinguishes between discovery motions and “merits” motions for purposes of access).

*United States v. Valenti*, 987 F.2d 708 (11th Cir. 1993) (affirms authority of district court to seal bench conference and articulate reasons later to satisfy a compelling interest; holds “dual-docketing system” unconstitutional).

Rule 5(d), Federal Rules of Civil Procedure (discovery materials not to be filed until “used in the proceeding” or until the court orders filing).

Rule 79, Federal Rules of Civil Procedure (requires maintenance of “civil docket” and chronological listing of entries thereon).

Rule 412(c), Federal Rules of Evidence (provides for sealing of documents and *in camera* hearing on requests to use sexual conduct evidence).

United States District Court, District of New Jersey, L. Civ. R. 5.3 (establishes comprehensive procedures for electronic filing, including motions to seal).

United States District Court, District of South Carolina, Local Civil Rule 5.03 (establishes procedure for motions to seal).

*A full bibliography, updated periodically, may be found on The Sedona Conference® web site at [www.thesedonaconference.org](http://www.thesedonaconference.org)*

## CHAPTER 3 PROCEEDINGS IN OPEN COURT

### **Principle 1    The public has a qualified right of access to trials that can only be overcome in compelling circumstances.**

Public access to trials<sup>4</sup> on the merits reflects a long tradition in the United States. Trials have long been considered open to the public. Public access to trials is essential to the monitoring and oversight of the judicial process. Public access also allows the public to share the “communal” experience of the trial. Constitutional and common law rights of access compel the conclusion that trials are at the heart of the right of public access both for historical reasons and to vindicate concerns for the legitimacy and accountability of the judicial system.

Legitimate interests exist that may justify narrow restrictions on public access to trials. For example, governmental interests exist in protecting certain types of information, such as those pertaining to classified national security information, undercover operations, and confidential informants. Privacy interests, such as those related to juveniles and to sensitive medical information, may justify limited trial closure in some contexts. Property interests also exist which may warrant restrictions, such as those related to trade secrets. Moreover, the judicial system itself may, as an institution, require a limited restriction on immediate public access for purposes of, for example, “sidebar” conferences.

The United States Supreme Court has established a right of public access to criminal trials derived from the First Amendment. The Court has also established standards that govern any restriction of public access to criminal trials. See, e.g., *Press-Enterprise Co. v. Superior Court*, 464 U.S. 501, 508-510 (1984); *Globe Newspaper Co. v. Superior Court*, 457 U.S. 596, 606-607 (1982). Similar considerations apply to civil trials. A party seeking to restrict public access to a civil trial should demonstrate a substantial likelihood that a compelling interest will be prejudiced by allowing public access and that no alternative other than closure can adequately protect the threatened interest. Any restriction on public access ordered by the court should be narrowly tailored. The trial court should also make findings of fact and conclusions of law adequate to justify the closure.

#### Best Practices

1. Closure of a civil trial on the merits is extraordinary and should be permitted only in rare circumstances where compelling interests leave no alternative.
2. Any closure must be no broader than absolutely necessary, and should be strictly limited to that portion of the trial that requires closure.
3. Alternatives to complete closure should be employed whenever possible. For example, counsel and witnesses may be directed to avoid references to particular facts or subjects (if such a direction will not deprive a party of the right to a fair trial) or, if a witness has a legitimate privacy interest justifying protection of his or her identity, testimony may be taken anonymously.
4. If closure is necessary, the court should consider providing access through other means (such as providing the public with a prompt transcript) if this can be done without compromising the compelling interest that required closure.
5. Courts should require parties who contend that closure of a portion of a trial will be necessary to raise the issue through a written motion in advance of trial, to allow the court ample time for consideration, and to permit the full exploration of alternatives to closure (such as orders precluding references to certain matters and/or providing for substitutes for evidence that implicates confidentiality concerns). Such a closure motion should be heard at a hearing duly noticed and open to the public. The hearing on the motion should be recorded or transcribed.

<sup>4</sup> As used herein, “trials” encompasses jury and nonjury trials as well as any judicial proceeding in court or on the record, except those conducted *in camera*.

6. Any order of closure should be based on a complete statement of the reasons for closure and of the findings of fact that support the closure order in sufficient detail to allow appellate review. The order should include express findings identifying the compelling interest that requires closure and the reasons why less restrictive alternatives are insufficient.
7. If the interest that led to closure loses its compelling importance with the passage of time, the transcript of the closed portion of the trial should be made available to the public.
8. The need to hold certain brief discussions during the course of a trial outside the hearing of the jury under circumstances where it is impractical to excuse the jury from the courtroom justifies the use of “sidebar” conferences that are inaudible to the jury and the public. Because such conferences involve the discussion and resolution of procedural and substantive issues integral to the conduct of the trial, however, these should remain subject to the right of public access and, absent compelling justification, transcripts of such proceedings should be promptly made available to the public (subject to whatever protections are necessary to keep matters from coming to the attention of the jury).
9. Conferences in chambers are occasionally used to address issues that arise in the course of a trial. Because it is not practical to admit the public to in-chambers conferences, the use of such conferences to address the merits of procedural and substantive issues (as opposed to, for example, scheduling matters or other routine discussions of no genuine public interest) should be avoided unless there are compelling circumstances that justify exclusion of the public (for example, where the subject under discussion involves a jury issue that implicates protected privacy rights of a juror). If procedural or substantive issues are discussed in chambers, the proceedings should be recorded and transcripts promptly made available to the public (unless, again, compelling circumstances justify confidentiality).
10. In-chambers proceedings should never be used to prevent public access to trial proceedings that could not be closed to the public if they took place in a courtroom.
11. Any closed trial proceedings must be transcribed or recorded in the same manner as open proceedings, so that access may be provided if the closure order is later reversed, or if the interests that require closure are later waived or no longer require protection.

#### Examples

1. On the first day of trial, the attorney for the plaintiff makes an oral application to the court to seal the trial, contending that his client might be embarrassed by public disclosure of “private facts.” No notice of this motion appears on the public docket. The attorney presents no facts to support the likelihood, nature, or extent of the damage his client would suffer by the public disclosure of the “private facts,” and the attorney does not present any reasons why a less restrictive alternative to sealing the trial would not satisfy the concern (for instance, sealing only particular testimony or evidence). Based on these deficiencies, the court denies the application.
2. Consistent with the local rules of the court, a plaintiff files a motion to seal expert testimony that will be offered by both parties in a patent infringement trial. The experts will testify on the damages sought by the plaintiff, and their testimony will be premised on financial data held confidentially by the plaintiff. The plaintiff submits an affidavit with its motion, which explains the nature of the data and why it is confidential. The affidavit also explains the competitive harm that will be visited on the plaintiff if the data becomes public. The court hears the motion, makes findings of fact and conclusions of law, and holds that specified portions of the experts’ testimony revealing the financial data will be sealed. The court’s ruling is filed for public review.

**Principle 2 The public has a qualified right of access to the jury selection process.**

Public access to the jury selection process promotes fairness by allowing the public to verify the impartiality of jurors, who are key participants in the administration of justice. Moreover, public access enhances public confidence in the outcome of a trial because public access assures those who are not attending that others may observe the trial. Public access also vindicates the societal concern that wrongdoers are brought to justice by individuals who are fairly selected to be jurors. Consistent with these interests, courts should not conceal from the public information that might bear on the ability of jurors to decide the matter before them impartially. Public access fosters discussion of government affairs by protecting the full and free flow of information to the public.

There may be circumstances in which some restriction to public access is necessary to ensure the safety or well-being of individual jurors, or to address personal privacy concerns.<sup>5</sup> In certain civil cases, courts may order that access to juror identities be limited or deferred to protect the safety of jurors, or to ensure their verdict is a product of the evidence admitted at trial rather than outside interference. See *Richmond Newspapers, Inc. v. Virginia*, 448 U.S. 555 (1980); *In re Globe Newspaper*, 920 F.2d at 94; *In re Baltimore Sun Co.*, 841 F.2d 74 (4th Cir. 1988); *Sullivan v. National Football League*, 839 F. Supp. 6 (D. Mass 1993).

Various interests are cited in opposition to public access to the juror selection process. These interests include juror privacy, ensuring that an adequate number of persons are willing to serve as jurors, protection of the integrity of the jury system and avoiding a circus-like atmosphere. But the difficulty in attracting qualified jurors, or the fact that a trial may be the subject of intense media coverage, are insufficient reasons to deny public access to the jury selection process. These reasons would render the right of access meaningless—the very demand for openness would defeat its availability. Thus, personal preferences of jurors, a judge's unwillingness to expose jurors to press interviews, or a judge's concern that jurors may disclose what transpired during deliberations do not, by themselves, warrant anonymous juries or restrictions on public access.

The qualified right of public access to the jury selection process does not extend to the deliberations of jurors, which traditionally occur in secret. In the case of jury deliberations, that secrecy is reinforced by substantive evidentiary rules that prevent jury verdicts from being impeached by testimony concerning the jury's internal deliberations in most instances.

However, in cases in which the conduct (or misconduct) of the jury itself becomes an issue that is the subject of testimony and/or other proceedings before a court, such proceedings, like other trial proceedings, are subject to the right of public access and should remain confidential only if compelling reasons (such as legitimate interests in juror privacy or in protecting a criminal investigation) justify confidentiality. Concealing a juror's misconduct is not by itself a legitimate privacy concern. While a court may take steps to prevent remaining jurors from being tainted by such proceedings, that is not in itself a reason for denying public access; rather, the steps to be taken should be similar to those used by the court to prevent jurors from having access to other possibly prejudicial information about a case (*i.e.*, instructions to avoid news coverage or, in some cases, sequestration). See *United States v. Edwards*, 823 F. 2d 111 (5th Cir. 1987).

Courts may limit an attorney's or party litigant's ability to interview jurors regarding their verdict or deliberations, or may require a showing of good cause before allowing such post-verdict interviews. Such orders do not themselves implicate the public's right of access to any public information, but only limit the behavior of lawyers and litigants. However, courts ordinarily should not limit the public's ability to interview jurors after the conclusion of a trial. Compare *In re Baltimore Sun Co.*, 841 F.2d at 75-76; *State v. Neulander*, 801 A.2d 255 (N.J. 2002), and *In re Express News Corp.*, 695 F.2d 807, 810 (5th Cir. 1982); *United States v. Antar*, 38 F.3d 1348, 1364 (3<sup>rd</sup> Cir. 1994).

---

5 Proposed Fed. R. Civ. P. 5.2 would mandate that certain personal identifiers be redacted from documents filed with the United States Courts. Neither the text of the proposed rule nor the Advisory Committee Notes explicitly address documents created or filed by the courts themselves, such as juror information. While redaction of personal identifiers from juror information may be considered consistent with the overall intent of the proposed rule, redaction of such information by court personnel is inconsistent with the intent of the Advisory Committee to place such duties on litigants and counsel.



## Best Practices

1. Empanelling an anonymous jury or closing jury *voir dire* (“jury secrecy” procedures) are extraordinary and should be undertaken only in rare circumstances where exceptionally important interests leave a trial court with no practical alternative. Although such circumstances have sometimes been found to exist in criminal cases (especially ones involving organized crime), it would be extremely rare for such circumstances to be present in a civil case. If initial questioning of jurors is conducted through written questionnaires, these should be available to the public.
2. Before answering oral or written questions from the court or the parties, potential jurors should be informed about whether and how the court will protect the confidentiality of any information provided. Such procedures help alleviate concerns that the government will compel them to disclose personal or confidential information without adequate privacy protections and thereby ensure that sufficient numbers of citizens will agree to serve and will not avoid service because of privacy concerns.
3. Any jury secrecy order should be no broader than absolutely necessary, should be strictly limited to highly personal juror information that requires protection, and should be entered only on the affirmative request of an individual juror. Alternatives to jury secrecy should be employed whenever possible.
4. If jury secrecy is necessary, a trial court should provide access through other means (*i.e.*, a transcript) if this could be done without compromising the overriding interest that required secrecy in the first place.
5. If the interest that led to jury secrecy loses its overriding importance with the passage of time, a transcript of the closed portion of the jury proceeding or the names of anonymous jurors could be made available to the public upon application to the court by an interested party. Stronger reasons to withhold juror names and addresses (such as jury tampering) normally arise during trial rather than after a verdict is rendered.
6. Trial courts should require parties who anticipate that jury secrecy may be necessary to raise the issue through written motions, to allow a trial court ample time for consideration and to permit the full exploration of alternatives to secrecy (such as change of venue). Such motions should be heard at a hearing duly noticed and open to the public.
7. Courts should freely allow nonparties who oppose jury secrecy to submit papers and make arguments addressing any secrecy motion. Intervention should be liberally granted for this purpose.
8. Any order of jury secrecy should be based on findings of fact that support the secrecy order, including express findings identifying the compelling interests that require secrecy and the reasons why less restrictive alternatives are insufficient.
9. Jury secrecy orders, while appearing to be interlocutory in nature, should be appealable by mandamus or under the collateral order doctrine.
10. Any closed jury proceeding should be recorded and transcribed, so that access may be provided if the secrecy order is later reversed, or if the interests that require secrecy are later waived or no longer require protection.

## Examples

1. In a civil action brought by a government agency against a well-known entertainer, the defendant approaches the court on the first day of trial and requests that the *voir dire* be sealed. The defendant’s argument is that juror candor in answering questions will be compromised by

the attendance of the press. In deciding the motion, the court considers if adequate public notice was given that the request would be made. Second, the court considers the proposition that candor might be compromised, noting that the application is solely premised on the defendant's celebrity status and press coverage. In the absence of a showing of facts supporting the proposition, the court denies the motion.

2. During the jury selection process in a civil action arising out of the sexual abuse of minors, the court advises the prospective jurors that they will be asked during *voir dire* if they ever experienced or witnessed sexual abuse. Given the nature of such questions and the (presumably) private nature of "yes" answers, the court gives jurors an opportunity to answer at sidebar in the presence of counsel. The sidebar conferences are transcribed. Two jurors state that parents or other relatives sexually abused them. One of these jurors testified about the abuse at a criminal trial. The other never reported the abuse. The judge releases the transcript of the *voir dire* of the first prospective juror but, after determining that the second juror would suffer psychological harm if the abuse became public, seals the transcript as to her. The judge puts findings of fact and conclusions of law on the record to justify the sealing.

### **Principle 3 Absent a compelling interest, the public should have access to trial exhibits.**

The right of public access to trial proceedings includes the right of public access to evidence admitted during a trial, including the testimony that is memorialized in the transcript. *See Nixon v. Warner Communications*, 435 U.S. 589, 609 (1978). Admitted evidence should be fully available to the public on a contemporaneous basis, and the standard for sealing evidence should be the same as that for closing a courtroom: That is, only compelling interests may justify sealing, and any order denying access must be based on findings of fact and conclusions of law demonstrating such an interest.<sup>6</sup>

It must be recognized, however, that logistical problems may foreclose contemporaneous access to trial exhibits in particular cases. This may simply be a question of practicality. *See, e.g., In re Application of National Broadcasting Co.*, 635 F.2d 945 (2d Cir. 1980). In such circumstances, the court, the parties, and the person seeking access should confer in an attempt to arrive at a procedure acceptable to all. When circumstances warrant, the public interest may be satisfied by providing access to the trial proceedings when the exhibit is admitted, rather than access to the exhibit itself. When access is sought to evidence introduced through means of novel technology such as computer generated or enhanced imagery, the court may be vested with wider discretion in deciding whether and how access may be allowed. *See In re Providence Journal Co.*, 293 F.3d 1 (1st Cir. 2002). Moreover, public access to trial exhibits is inhibited by a prevailing practice in most American trial courts to return trial exhibits to the parties after the time for filing an appeal has expired, or an appeal has been taken and resolved. Because the physical exhibits are not maintained as public records, these are no longer subject to enforceable public access rights. *See Littlejohn v. Bic Corp.*, 851 F.2d 683 (3rd Cir. 1988).

#### **Best Practices**

1. Courts should reasonably and promptly accommodate requests for access to exhibits admitted at trial and not under seal.
2. Providing access to trial evidence so that copies, recordings and/or photographs can be made of the evidence should be routine. Access should be denied only in rare circumstances where compelling interests leave the court with no practical alternative. Any denial of access to trial exhibits should be no broader than absolutely necessary and should be strictly limited to evidence that requires protection. Alternatives to denial of access should be employed whenever possible.

<sup>6</sup> The Working Group reached no consensus on a right of public access to excluded evidence. On the one hand, such evidence, by its very exclusion, has been deemed by the judge to be irrelevant to the jury function. On the other hand, access to excluded evidence may allow the evaluation of the judge's role as "gatekeeper" and the overall fairness of the trial. In addition, evidence marked during trial becomes part of the record on appeal, even if excluded, indicating that appellate courts consider excluded evidence to be part of the adjudicative process.

3. A party that intends to request that the court seal evidence it expects to be admitted at trial should make its request by written motion before trial. The court should hear and decide motions to seal admitted trial exhibits after other parties have had time to oppose the request, or non-parties have had time to request leave to intervene to oppose the request. Absent the most exigent circumstances, trial courts should deny any request for denial of access that is not made in time to allow such notice.
4. The hearing on the motion should be recorded.
5. If denial of access is necessary, a trial court should consider providing access through other means (i.e., providing access at the conclusion of the trial) if this can be done without compromising the overriding interest that required denial of access. If the interest that led to denial of access loses its overriding importance with the passage of time, access to trial evidence should be granted at the earliest possible time.
6. Any order denying access to evidence should be based on a complete statement of the reasons for denial and of the findings of fact that support the order denying access, including express findings identifying the compelling interests that require denial of access and the reasons why less restrictive alternatives are insufficient.

#### Examples

1. A plaintiff receives documents in discovery from the defendant in a product liability case. Prior to trial, the documents are marked confidential pursuant to a protective order entered in the case, and the plaintiff is prohibited from disseminating them publicly. Plaintiff identifies the documents on an exhibit list exchanged before trial, and during the trial these are marked and admitted in evidence. At no time does the defendant move to have the trial exhibit sealed. The exhibits are available to the public.
2. Same facts as above, except that the trial has ended and the exhibits have been returned to the parties who introduced the exhibits, and the press seeks to obtain copies of the admitted exhibits from the plaintiff. Although the protective order once protected the documents from disclosure, the plaintiffs' attorney provides the exhibits to the press because they were admitted in open court. The defendant learns of the press's request and seeks to enforce the protective order to bar the plaintiff from providing the exhibits to the press. The court rejects the defendant's effort to invoke the protective order because the documents, having been received in evidence in open court, are no longer properly subject to protection.
3. A plaintiff receives documents in discovery from the defendant in a product liability case. The documents are marked "confidential" pursuant to a protective order entered in the case and the plaintiff is prohibited from disseminating them publicly. Plaintiff identifies the documents on an exhibit list exchanged before trial and the defendant requests that the documents only be admitted under seal. The defendant moves before trial to seal the exhibits on the basis that the exhibits contain proprietary information and trade secrets. The court holds a hearing and grants the motion, protecting genuinely proprietary information from disclosure. The court admits exhibits for which no compelling need for protection has been established.

## CHAPTER 3

### Selected Bibliography

In re *Providence Journal Co.*, 293 F.3d 1 (1st Cir. 2002) (right of access to attend criminal trial and pretrial proceedings extends to “documents and kindred materials;” defendants’ right to fair trial constitutes compelling interest sufficient to allow restriction of access; common law, but not First Amendment, right of access held to encompass duplication of evidence, but where “cutting-edge technology” in issue, trial court given discretion to accommodate access).

*ABC, Inc. v. Stewart*, 360 F.3d 90 (2d Cir. 2004) (vacating in part order closing *voir dire* examination of potential jurors; recognizes that after-the-fact release of transcript no substitute for presence; trial court failed to demonstrate interest in assuring juror candor sufficient to seal and failed to use available alternatives to sealing).

*Publicker Indus., Inc. v. Cohen*, 733 F.2d 1059 (3d Cir. 1984) (First Amendment and common law rights of access extended to civil proceedings).

In re *Application of National Broadcasting Co.*, 635 F.2d 945 (2d Cir. 1980) (common law right of access to inspect and copy judicial records extends to evidence introduced at trial, whether documentary or of other nature, under reasonable procedures to be determined by the court).

*United States v. Antar*, 38 F.3d 1348 (3d Cir. 1994) (presumptive right of access applies to *voir dire* of examination of potential jurors; concludes that trial court erred in sealing transcript without adequate notice and findings of fact; modifies trial court restrictions on post verdict interviews of jurors by press).

*United States v. Simone*, 14 F.3d 833 (3d Cir. 1994) (presumptive right of access applies to post verdict examination of jurors into possible misconduct; strikes down closure of examination given failure of trial court to articulate “overriding interest;” after-the-fact receipt of transcripts not equivalent to actual presence).

*Little John v. BIC Corp.*, 851 F.2d 673 (3d Cir. 1988) (common law presumption of access applies to documents initially produced in discovery pursuant to protective order and later admitted into evidence at trial, but exhibits returned to party after trial are no longer judicial records and disclosure cannot be compelled).

In re *Perrigo Co.*, 128 F.3d 430 (6th Cir. 1997) (injunction prohibiting magazine from publishing materials filed under seal violated First Amendment; umbrella protective order pursuant to which documents filed under seal without good cause determination invalid).

*United States v. McDougal*, 103 F.3d 651 (8th Cir. 1996) (affirms refusal to allow media access to videotaped depositions of President Clinton, although introduced at criminal trial and transcript released; concludes that, under circumstances presented, videotape itself not a “judicial record;” rejects “strong” presumption of access recognized by other circuits and defers to sound discretion of trial court).

*United States v. McVeigh*, 119 F.3d 806 (10th Cir. 1997) (First Amendment right of access does not extend to suppressed evidence or evidence inadmissible at trial).

*Goff v. Graves*, 362 F.3d 543 (8th Cir. 2004) (affirms sealing of depositions of confidential prison informants in §1983 action; preservation of institutional security and protecting against retaliation are compelling interests to issue protective order and to seal portion of record).

28 C.F.R. Sec. 509 (sets out U.S. Department of Justice policy with regard to open judicial proceedings, civil and criminal).

*A full bibliography, updated periodically, may be found on The Sedona Conference web site at [www.thesedonaconference.org](http://www.thesedonaconference.org)*

## CHAPTER 4 SETTLEMENTS

**Principle 1** **There is no presumption in favor of public access to unfiled settlements, but the parties' ability to seal settlement information filed with the court may be restricted, due to the presumptively public nature of court filings in civil litigation.**

A dichotomy currently exists between open courts and private alternative dispute resolution such as arbitration and mediation. In many cases, parties may be willing or able to waive litigation in favor of such private dispute resolution, where the confidentiality of both the process and its outcome can be contractually assured. Similarly, litigants possess broad discretion to contract privately for confidentiality as a condition to settling even litigated disputes. In such cases, confidentiality, like other settlement terms, becomes a matter of private agreement to be enforced pursuant to applicable contract law.

There is a strong public policy in favor of settlement. Confidentiality of settlement terms is generally believed to encourage such settlements, and in the majority of cases, the parties need not make their settlement public by filing it with the court. Courts will generally enforce private confidentiality agreements so long as they merely restrict the voluntary disclosure of information and do not prohibit disclosures required by law or court order. Because the agreements can be reached without any judicial involvement, and the settlement itself is rarely filed with the court, these confidential settlements do not implicate any right of public access. See *Gambale v. Deutsche Bank AG*, 377 F.3d 133, 143 (2d Cir. 2004); *Herrnreiter v. Chicago Housing Authority*, 281 F.3d 634, 636-37 (7th Cir. 2002); *Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 781, 788-89 (3d Cir. 1994); Laurie K. Doré, *Settlement, Secrecy, and Judicial Discretion: South Carolina's New Rules Governing the Sealing of Settlements*, 55 S.C. L. Rev. 791, 799-800 (2004).

In many cases, however, a confidential alternative to litigation may not be available to disputing parties. Likewise, settling parties may not wish to rely solely upon private agreement to ensure confidentiality and may choose instead to more deeply involve the court in their confidential compromise. See Laurie K. Doré, *Settlement, Secrecy, and Judicial Discretion: South Carolina's New Rules Governing the Sealing of Settlements*, 55 S.C. L. Rev. 791, 801-04 (2004). In utilizing a public court to resolve their dispute or enforce its confidential settlement, the litigants invoke the jurisdiction of a forum that is subject to public oversight and monitoring. As such, confidentiality may no longer be a matter within the exclusive control of the parties. Confidentiality agreements between the parties regarding settlement do not bind the court, and if the parties wish to file their settlement agreement under seal, the court must exercise independent judgment and comply with applicable legislative and judicial standards<sup>7</sup> before issuing any sealing order incident to a settlement. See *Gambale v. Deutsche Bank AG*, 377 F.3d 133, 139-42 (2d Cir. 2004).

Although contract doctrine in some jurisdictions may invalidate confidentiality clauses that are illegal, unconscionable, or contrary to public policy, a strong and well-established public policy favors alternative dispute resolution (ADR) and the private settlement of disputes. Thus, notwithstanding the potential public interest in the resolution of disputes involving statutory or public rights (*i.e.*, consumer or employee claims) or public health and safety, little public oversight of confidentiality in alternative dispute resolution or private settlement currently exists. Attorneys should thus ensure that their clients are fully aware of all available public and private dispute resolution processes and advised that confidentiality cannot be expected or assured in a public forum.

### Best Practices

1. At or before the commencement of litigation, attorneys should confer with their clients to determine whether private dispute resolution is available and would be preferable to traditional

---

<sup>7</sup> See Chapter 4, Principle 2 below.

litigation in the courts. Among other things, attorneys should present these options to clients in the context of the clients' needs or desires to maintain confidentiality.

2. Attorney should discuss with clients that certain disputes present "classic" matters for private resolution, such as breach of a commercial contract between business entities. In contrast, other disputes, such as those brought by individual consumers or employees to vindicate statutory rights or involving public entities or officials, may not be appropriate for private dispute resolution given the public interest in their resolution.
3. In discussing the above options with clients, attorneys should also discuss the available mechanisms for enforcement of any breach of confidentiality by adversaries.

#### Examples

1. Two parties to a commercial agreement include in that agreement a provision for mandatory arbitration or mediation. The agreement provides for a sale of goods by one party to the other. No public health or safety concerns are implicated. Under these circumstances, there is no need for public oversight of the dispute resolution process and the parties may ensure confidentiality through contract. Enforcement of any settlement would be through contract law principles.
2. An individual receives telecommunications services from a large business entity. The consumer, under the terms of a standard agreement that was mailed to him, is required to arbitrate any future dispute that he has with the business entity. Under the terms of a confidentiality provision in the arbitration agreement, no public access is available to the facts or nature of the dispute, the arbitration proceedings, the terms of any award or settlement, or the services rendered by the arbitrator in this particular case. This dispute, which may implicate the rights of similarly situated consumers, involves statutory claims that have both remedial and deterrent objectives and that arguably implicate a public interest broader than the immediate parties. The confidentiality of this alternative forum raises policy concerns absent from example 1 and not addressed by the Working Group.
3. A business entity intends to commence an antitrust action against a competitor. Rather than proceed to litigation, the parties agree to arbitrate or mediate. A settlement is reached under which the competitors agree to divide markets on a geographic basis, in violation of antitrust laws. A startup company formed by former employees with knowledge of the confidential settlement agreement files suit, alleging that the settlement violates antitrust laws. The settling parties should not have any expectation that the confidentiality of the settlement will be maintained, inasmuch as the legality of the settlement itself has been questioned and is relevant to the action.

**Principle 2    Settlements filed with the court should not be sealed unless the court makes a particularized finding that sufficient cause exists to overcome the presumption of public access to judicial records.**

Parties that resort to a public forum and that enter into a settlement thereafter may have legitimate interests that warrant confidentiality. However, given the public right of access to public forums, sealed settlements should be the exception and not the norm. Courts should assure that settlements that are filed with the court are not sealed unless good cause exists; and unless specific findings of fact and conclusions of law are made and are available for public review. Attorneys who choose to file a settlement with the court should not seek to seal that settlement unless they are satisfied that such a showing can be made.

Courts are public forums. As discussed in Chapter 2, there is a presumption of access to courts and to information filed with courts, including settlement agreements. Thus, information considered to be confidential by parties and filed with courts may, by the act of filing, become public records subject to public access. *Jessup v. Luther*, 277 F.3d 926 (7th Cir. 2002); *Bank of America Nat'l Trust & Savings Ass'n v. Hotel Rittenhouse Assoc.*, 800 F.3d 339 (3d Cir. 1986). This presumption of

public access, however, is arguably weaker (and thus more easily rebutted) for “settlement facts” that relate to the specific terms, amounts, and conditions of a settlement involving non-governmental, private litigants than the presumption that attaches to information more central to the adjudicatory function of courts. *See Gambale v. Deutsche Bank AG*, 377 F.3d 133, 143-44 (2d Cir. 2004) (distinguishing between settlement amount and summary judgment documents). In considering whether to seal a settlement, then, a court might appropriately distinguish between settlement information that would not exist but for the settlement and “adjudicative” facts that may be relevant to the underlying merits of the settled controversy.

Parties may have legitimate interests in the confidentiality of all or part of a settlement. Parties may also have justifiably relied on a promise of confidentiality in entering into a settlement. At the same time, however, a sealed settlement may affect the interests of the general public and collateral litigants. To overcome the presumption of public access, then, parties must establish sufficient cause and satisfy applicable tests established by legislatures and courts to govern the sealing of a settlement, in whole or in part. In addition, the judge should make specific findings of fact and conclusions of law on any application to seal a settlement to determine whether the presumption of public access has been overcome.

Despite resort to a public forum, the parties may elect to avoid any question of access to the terms of a settlement by choosing not to file it. For example, parties may enter into a settlement agreement and then file a voluntary dismissal under Fed. R. Civ. P. 41(a)(1). Such a dismissal requires no judicial action and the settlement agreement would not be submitted to the court. Alternatively, a party could move for dismissal under Fed. R. Civ. P. 41(a)(2), the resolution of which does not require approval of any settlement agreement. Under either procedure (or their state equivalents), the settlement agreement does not become a public record. The unfiled settlement will not trigger any presumption of public access and will instead have the status of any other private contract. *Compare SmithKline Beecham Corp. v. Pentuch Group, P.L.C.*, 261 F. Supp. 2d 1002, 1004-08 (N.D. Ill. 2003) (Posner, C.J., sitting by designation) (alleged illegality of settlement agreement not subject to review under either procedure) *with Fomby-Denson v. Department of the Army*, 247 F.3d 1366, 1374-75 (Fed. Cir. 2001) (declining *sua sponte* to enforce confidential settlement agreement as contrary to public policy).

Settlements that are both filed and sealed appear to be infrequent, at least in federal courts. R. T. Reagan, *et al.*, *Sealed Settlement Agreements in Federal District Court* (Federal Judicial Center, 2004). Confidential settlements, however, are more common and of broader concern. Parties are largely free to agree between themselves to confidentiality provisions in settlements. If the parties enlist the court’s assistance concerning the settlement, however, the court should independently scrutinize any confidentiality provision. For example, the parties may request that the court retain enforcement jurisdiction to oversee the fulfillment of the settlement terms, or they may file a separate action for specific performance of the settlement or to recover damages for its breach. In such cases, the court should not enforce any non-disclosure or confidentiality provision that is not supported by a specific and sufficient showing of good cause.

#### Best Practices

1. Before attempting to seal a settlement, attorneys should confer with their clients to ensure that legitimate privacy, commercial or similar confidences exist that warrant confidentiality.
2. When negotiating the terms of a settlement, attorneys should confer among themselves with regard to any need for confidentiality and attempt to reach agreement on legitimate grounds for confidentiality.
3. Attorneys should not seek to seal settlements unless they are satisfied that grounds exist for a sealing order.
4. In considering whether to seal a settlement or enter a confidentiality order incident to a settlement, courts should distinguish between “settlement facts,” such as the amount, terms and

conditions of a compromise, and “adjudicative facts” that are relevant to the merits of the underlying controversy. The former, which arise out of the settlement process itself, might warrant a sealing order. Care should be taken in extending any such order to the latter so as to avoid suppressing information relevant to other cases, public health or safety, or other legitimate public interest.

5. In negotiating a confidential settlement agreement, attorneys should incorporate into any confidentiality provision an explicit exception for disclosures required by law or court order.

#### Examples

1. An individual plaintiff and a corporate defendant have entered into a settlement of a personal injury action. The defendant, as a matter of corporate practice, does not reveal the monetary amount of any settlements. The defendant insists, and the plaintiff agrees, on a confidential settlement. The parties do not contemplate filing the settlement with the court, as there is no basis for a sealing order.
2. An individual plaintiff and a corporate defendant have entered into a settlement of a personal injury action. The defendant settled to avoid the publication of internal documents at trial and on the express condition that the amount of the settlement would be confidential. The parties want the terms of the settlement embodied in an order by which the court retains jurisdiction to enforce the settlement. They move to seal the settlement. While the settlement itself would be presumed to be a public document if filed, the presumption of public access is weak as to the amount of the settlement. The court seals only the amount of the settlement.
3. An individual plaintiff who developed certain software is in litigation with a corporate defendant. The litigation arises out of the corporate defendant’s alleged failure to develop and market new applications arising out of the software. The parties enter into a settlement agreement. The terms of the settlement provide for the parties to share source codes of the defendant’s applications. Both parties, who are in a very competitive field of business, deem the source codes highly confidential. The parties agree that the settlement should be confidential. Neither party trusts the other and both contemplate injunctive relief and contempt should the source codes be misused. They agree to file a motion to seal the settlement. The source codes are described in detail so that there can be no misunderstanding of the scope of the settlement in any future enforcement action. The court issues an order sealing only that part of the settlement that reveals the source codes.

### **Principle 3 Settlement discussions between parties and judges should not be subject to public access.**

Courts primarily exist to resolve disputes. Disputes may be resolved in a number of ways, including settlement. A strong public policy supports settlement and the “just, speedy, and inexpensive determination of every action.” Fed. R. Civ. P. 1. Thus, judges should be expected to encourage settlement and to participate in settlement discussions.

Judges are public officials, however, and, as such, are subject to oversight and monitoring by the public. Thus, when a judge participates in settlement discussions between the parties or is otherwise “injected” into the settlement process, the judge’s actions are arguably subject to public monitoring and oversight. The desire or need for such oversight and monitoring may be heightened when settlement discussions affect public health and safety.

However, several factors argue against public access to settlement negotiations even when they may involve the court. First, settlement negotiations require candor, and public access might discourage a party from revealing information necessary for self-evaluation and compromise. Second, settlement discussions are often conducted on an *ex parte* basis, where information is exchanged with the judge for settlement purposes only and is never shared with the adversary. Third, and most significantly, in promoting settlement, the judge acts as a facilitator, rather than as an adjudicator.



Because the judge is not engaged in “decision-making,” the rationale for public oversight and monitoring is significantly diminished.

Indeed, in many cases, the parties may privately settle their dispute without filing their settlement or submitting it for approval or other action by the court. In these cases, the case is dismissed on stipulation. No judicial record exists and the judge has neither the need nor the power to approve or disapprove of the settlement. In such cases, where the judge has no approval role and serves merely as a mediator or facilitator for the parties’ private negotiations, any presumption of public access is weak, if not non-existent.

If public access is to be denied on this premise, however, the judge should take care not to step into a judicial role concerning the settlement. To protect the confidentiality of their settlement, the parties should not file their agreement with the court or seek judicial “approval” of their compromise. If parties voluntarily elect to file their settlement agreement in order to facilitate its subsequent enforcement, their action may create a judicial record, trigger a presumption of public access, and forfeit the confidentiality of the settlement.

#### Best Practices

1. A judge may act as an intermediary or facilitator in settlement negotiations between the parties to a case. Alternatively, the judge may refer the case for confidential, court-annexed alternative dispute resolution such as mediation. So long as the court does not step into an adjudicatory role and the settlement agreement is not filed with the court, no presumption of public access to the settlement discussions or to any settlement agreement will result.
2. Absent a statute or rule which requires otherwise, attorneys should not ask a judge to “approve” a settlement that they wish to keep confidential, file that settlement with the court or request that the terms of the confidential agreement be entered as orders of the court. A judge should not seek to approve a private settlement unless required or requested to do so.
3. In cases where judicial approval of a settlement is legally required (*e.g.*, class actions), or in cases where the parties seek the court’s *imprimatur* on their settlement so that it can be entered as a consent decree enforceable through injunction, contempt or summary judgment, the settlement must be filed and submitted to the court. In such cases, the settlement agreement becomes a presumptively public judicial record, and proceedings leading to its formal approval are subject to a qualified right of public access.
4. A judge should not *sua sponte* suggest to the parties that a settlement might be kept confidential. In a case pending in federal court, however, a judge might appropriately suggest, as an alternative, that the court retain jurisdiction to enforce a settlement.
5. If the terms of a settlement are presented to a judge, the judge may express concern about any term that might arguably be illegal, unethical or unenforceable. However, it may be difficult for a judge to independently “police” the provisions of a settlement in this manner, as there will be no adversarial development of any issue.

#### Examples

1. The parties to a commercial dispute appear before a judge for a settlement conference. The judge conducts the conference in chambers and engages in *ex parte* discussions with the parties in an attempt to facilitate a settlement. A settlement is reached. The terms of the settlement are not put on the record and the settlement agreement is never filed. The settlement is private and there is no right of public access.
2. The same facts as (1) above, but the settlement is submitted to the court as a stipulation, with a motion that it be adopted by the court as an order. If the judge grants the motion, the judge gives the settlement a public *imprimatur* and the settlement becomes a public record.

3. The same facts as (1) above, but the settlement includes a provision whereby the parties agree to divide their state into districts and not compete with each other in certain districts. The judge cautions the parties of the possible illegality of the settlement and refuses to approve or otherwise facilitate the settlement.

**Principle 4 Absent exceptional circumstances, settlements with public entities should not be confidential.**

Public entities and officials, whether at the federal, state or local level, are public actors. By definition, their actions affect the public, whom they represent. The public thus possesses a significant interest in the monitoring and oversight of public officials and entities, even in litigation. Public entities are generally subject to open public meeting and/or open public record laws. Such laws, which seek to facilitate public monitoring and government accountability, may require disclosure of settlements involving the government or other public entities. Thus, when a public entity enters into a settlement, no expectation of confidentiality should exist, whether or not the settlement is filed with the court.

For these reasons, there should be a strong presumption against the confidentiality of any settlement entered into with a public entity or of any information otherwise disclosable under a public records law. See *Pansy v. Borough of Stroudsburg*, 23 F.3d 772, 792 (3d Cir. 1994). A particularly strong presumption of public access exists with regard to any monies paid by public entities in settlement. Only exceptional circumstances (such as those involving intimate personal information, the privacy of minors, or law enforcement needs) should warrant the confidentiality of these types of settlements. See generally Laurie K. Doré, *Settlement, Secrecy, and Judicial Discretion: South Carolina's New Rules Governing the Sealing of Settlements*, 55 S.C. L. Rev. 791, 809-10 (2004).

**Best Practices**

1. An attorney representing an individual or entity in litigation against a public agency should, before entering into settlement negotiations with the agency, consult with his or her client to determine whether the client has any proprietary or privacy interest in the terms of the settlement for which protections should be sought under applicable public records law or a court order. Absent any such protection, the attorney should caution his or her client against any expectation of confidentiality.
2. An attorney representing a public agency should, in the course of settlement negotiations with an adversary, caution the adversary against expecting any confidentiality of a settlement agreement, absent specific exemptions in the public records laws or a court order.
3. In determining whether to seal a settlement of a matter involving a public entity or official, a court should carefully consider relevant federal or state law. On the one hand, judges should be hesitant to seal a settlement if the information would be otherwise disclosable under a federal or state freedom of information or open public records statute. On the other hand, information that would be exempt from disclosure under such a law or separate privacy-related statute might merit a confidentiality order.

**Examples**

1. A business entity sues a state agency for breach of contract. The action arises out of alleged delay damages incurred by the plaintiff after the defendant agency failed to accept goods on a certain date. The settlement agreed to by the parties includes, at plaintiff's insistence, a confidentiality provision. No legitimate basis for confidentiality exists.
2. An individual sues a state agency for wrongful disclosure of her private medical information. The defendant agency admits that it erred in disclosing the information. The parties enter into a settlement which, at plaintiffs' insistence, seals all facts relevant to the suit, including plaintiffs' medical information. The plaintiffs' information may be sufficiently confidential to justify sealing the settlement or issuing a confidentiality order.

**Principle 5 An attorney’s professional responsibilities may affect considerations of confidentiality in settlement agreements.**

The obligation of an attorney to maintain a client’s confidences is fundamental to the attorney-client relationship. An attorney must thus take steps to arrive at a settlement that protects a client’s confidential information. Consistent with this obligation, an attorney must take client confidences into account during settlement negotiations and may seek an agreement to limit the voluntary disclosure of confidential information as a condition of settlement.

In all settlement negotiations, however, an attorney should consider and adhere to all applicable standards of professional responsibility. For example, certain nondisclosure provisions may violate ethical rules that prohibit restrictions on another attorney’s practice of law. The ethical rules of a jurisdiction may similarly prohibit a settlement that purports to restrict an attorney from using information gained in one case in other related cases. Additionally, a confidential settlement should not prohibit the disclosure of information required by law or court order.

An attorney should also recognize that confidential information may concern public health and safety or may affect specific individuals (such as collateral litigants). Depending upon the ethical rules of a jurisdiction, attorneys may have limited discretion to reveal confidences when death, serious bodily injury or financial harm is imminent. *See Model Rules of Professional Responsibility*, Rule 1.6(b). Little guidance exists, however, concerning the factors that an attorney should consider in deciding to exercise this discretion. *See id.*, Rule 1.6, Comment 6. Moreover, unless otherwise required (*See, e.g.*, 17 C.F.R. 205, “Standards of Professional Conduct for Attorneys Appearing and Practicing Before the [Securities and Exchange] Commission in the Representation of an Issuer”), no mandatory duty to disclose exists.

**Best Practices**

1. Attorneys should familiarize themselves with applicable ethical rules and substantive law to determine what limitations exist on their negotiating confidential settlements that might include unethical, illegal, or otherwise unenforceable terms.
2. Attorneys should familiarize themselves with applicable ethical rules and substantive law to determine which circumstances may permit disclosure of otherwise confidential information.
3. Regardless of whether ethical rules prohibit a nondisclosure provision or whether an attorney has discretion to disclose a confidence, an attorney should discuss and attempt to resolve any concerns concerning confidentiality with his client.

**Example**

1. The parties to a products liability action are engaged in settlement negotiations. The product at issue is a widely distributed and well-known kitchen appliance. Through study of the defendant’s highly confidential design documents, obtained during discovery under a protective order, the plaintiff learns of the existence of a design defect in the product’s control panel that might cause a fire like that in plaintiff’s case. Plaintiff’s attorney knows of at least four other cases involving fires in the appliance. The defendant insists that it will not settle without a confidentiality agreement. The attorney confers with his client about the proposed settlement and the defendant’s confidentiality demand. The client decides to agree to the settlement and confidentiality demand. Absent an ethical rule or substantive law to the contrary in the jurisdiction, the information would be considered a “client confidence.” The attorney may not voluntarily reveal any information covered by the confidentiality clauses of the settlement agreement.

## CHAPTER 4

### Selected Bibliography

28 C.F.R. §50.23 (policy of the United States Department of Justice with regard to confidential settlements)

Eastern District of Michigan, Local Civil Rule 5.4 (placing temporal limit on sealing of settlement agreements)

District of South Carolina, Local Civil Rule 5.03 (establishing procedure for motions to seal and barring sealed settlements)st

*Poliquin v. Garden Way, Inc.*, 989 F.2d 527 (1st Cir. 1993) (recognizing court's inherent power to modify confidentiality order even after settlement)

*Gambale v. Deutsche Bank*, 377 F.3d 133 (2d Cir. 2004) (recognizing "weak" presumption of access to settlement amount given reliance on confidentiality and lack of public interest concerning amount).

*Pansy v. Borough of Stroudsburg*, 23 F.3d 772 (3d Cir. 1994) (holding that because settlement agreement between police chief and borough was not filed with, interpreted or enforced by the district court, it was not a judicial record subject to right of access doctrine: "Simply because a court has entered a confidentiality order over documents does not automatically convert those documents into 'judicial records' accessible under the right of access doctrine." However, court allowed newspaper to intervene to challenge propriety of confidentiality order under state Right to Know Act and Federal Rule 26 and required that "good cause" support any confidentiality order).

*Zurich American Ins. Co. v. Rite Aid Corp.*, 2004 345 F.Supp.2d 497 (E.D. Pa. 2004) (vacating an order sealing an arbitration award in corporate governance case, after balancing public policy favoring disclosure against speculative concerns of party favoring secrecy).

*Ashcroft v. Conoco, Inc.*, 218 F.3d 288 (4th Cir. 2000) (addressing impropriety of sealing confidential settlement in environmental tort lawsuit involving contaminated drinking water)

*Ford v. City of Huntsville*, 242 F.3d 235 (5th Cir. 2001) (vacating confidentiality order concerning settlement between city and employee because trial court failed to consider effect of order on Texas Public Information Act)

*Goodyear Tire & Rubber Co. v. Chiles Power Supply, Inc.*, 332 F.3d 976 (6th Cir. 2003) (creating new federal common law privilege protecting "secrecy of matters discussed . . . during settlement negotiations")

*Herrnreiter v. Chicago Hous. Auth.*, 281 F.3d 634 (7th Cir. 2002) (defendant in employment discrimination proceeding gave up claim of confidentiality with respect to settlement agreement when it made the terms of that agreement a subject of litigation by filing a motion to implement the settlement).

*Lynch, Inc. v. Samatamson, Inc.*, 279 F.3d 487 (7th Cir. 2002) ("No one supposes that there is any impropriety in a judge's conducting settlement discussions off the record").

*Jessup v. Luther*, 277 F.3d 926 (7th Cir. 2002) (reversing district court's denial of newspaper's motion to unseal settlement agreement in §1983 action because filed settlement agreement was a judicial record subject to public right of access).

*Union Oil Company of California v. Leavell*, 220 F.3d 562 (7th Cir. 2000) (finding order sealing virtually all judicial documents in case file unjustified; when the parties "call on the courts, they must accept the openness that goes with subsidized dispute resolution by public (and publicly accountable) officials . . . only genuine trade secrets, or information within the scope of a requirement such as Fed. R. Crim. P. 6(e)(2) . . . may be held in long-term confidence").

*Foltz v. State Farm Mut. Auto. Ins. Co.*, 331 F.3d 1122 (9th Cir. 2003) (discussing right of collateral litigants to obtain discovery, motions, and judicial records filed under seal in settled case)

*Phillips v. General Motors Corp.*, 307 F.3d 1206 (9th Cir. 2002) (examining right of access in settled product liability case to information produced by manufacturer concerning total number and aggregate dollar amount of settlements in similar fuel-fed fire cases)

*Chicago Tribune Co. v. Bridgestone/ Firestone, Inc.*, 263 F.3d 1304 (11th Cir. 2001) (remanding for Firestone to demonstrate need for confidentiality of discovery filed under seal in settled case)

*EEOC v. Nat. Children's Ctr., Inc.*, 98 F.3d 1406 (D.C. Cir. 1996) (holding that trial court abused its discretion in sealing consent decree in EEOC action and remanding for further explanation of decision to seal portions of the record and determination of good cause to support protective order governing deposition transcripts).

Joseph F. Anderson, Jr., *Hidden from the Public by Order of the Court: The Case Against Government-Enforced Secrecy*, 55 S.C.L. Rev. 711 (2004).

A. F. Blakley, *To Squeal or Not to Squeal: Ethical Obligations of Officers of the Court in Procession of Information of Public Interest*, 34 Cumberland L. Rev. 65 (2003).

Laurie Kratky Doré, *Public Courts v. Private Justice: It's Time to Let Some Sun Shine in on Alternative Dispute Resolution*, 81 Chicago-Kent L. Rev. xxx (forthcoming 2006).

Laurie Kratky Doré, *Settlement, Secrecy, and Judicial Discretion: South Carolina's New Rules Governing the Sealing of Settlements*, 55 S.C.L. Rev. 791 (2004).

Laurie Kratky Doré, *Secrecy By Consent: The Use and Limits of Confidentiality in the Pursuit of Settlement*, 74 Notre Dame L. Rev. 283 (1999).

Robert. Timothy Reagan, et al., *Sealed Settlement Agreements in Federal District Court* (Federal Judicial Center, 2004).

James E. Rooks, Jr., *Settlements and Secrets: Is the Sunshine Chilly?*, 55 S.C.L. Rev. 859 (2004).

*A full bibliography, updated periodically, may be found on The Sedona Conference web site at [www.thesedonaconference.org](http://www.thesedonaconference.org)*

## CHAPTER 5

## PRIVACY AND PUBLIC ACCESS TO THE COURTS IN AN ELECTRONIC WORLD

**Introduction**

The growing use of electronic filing and imaging technology makes it possible for courts to offer broader public access to case files through remote electronic access. There is increasing awareness, however, of the implications such broad public access to case files, especially through the Internet, has for personal privacy, and proprietary business information. In the United States court community, many have suggested that case files – long presumed to be open for public inspection and copying unless sealed by court order – contain private or sensitive information, trade secrets, and proprietary information that should be protected from unlimited public disclosure and dissemination in the new electronic environment.<sup>8</sup> Others maintain that electronic case files should be treated the same as paper files in terms of public access and that existing court practices are adequate to protect privacy, confidentiality, and intellectual property interests.

**Potential Privacy, Confidentiality, and Proprietary Implications of Public Access to Electronic Case Files**

Before the advent of electronic case files, the right to “inspect and copy” court files depended on one’s physical presence at the courthouse. The inherent difficulty of obtaining and distributing paper case files effectively insulated litigants and third parties from any harm -- actual or perceived -- which could result from misuse of private or proprietary information provided in connection with a court proceeding. The Supreme Court has referred to this relative difficulty of gathering paper files as “practical obscurity.” See *United States Department of Justice v. Reporters Committee for Freedom of the Press*, 489 U.S. 749, 770-71, 780 (1989) (recognizing “the vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information”).

Case files may contain private, sensitive or proprietary information such as medical records, employment records, financial information, tax returns, Social Security numbers and other personal identifying information, as well as customer lists, business plans, research data, and other proprietary business information. Allowing access to case files through the Internet, depending on how it is accomplished, can make such information available easily and almost instantly to anyone who seeks it. Personal, sensitive, or proprietary information, unless sealed or otherwise protected from disclosure, can be downloaded, stored, printed, and distributed.

The information contained in court records, particularly the personal information, is highly valued by data-mining companies that cull public records and integrate public record data with other sources of data and sell the information for profit. Because there are few, if any, legal limitations on how public court records may be used by those who obtain them, data-mining companies are able to freely exploit the information in court records for commercial purposes as marketing information or “competitive intelligence.”

These circumstances place into conflict two important policy considerations. First, public court records must easily be available to allow for effective public monitoring of the judicial system; and second, private or sensitive information in court files that is not germane to the public oversight role may require protection from indiscriminate public disclosure.

---

<sup>8</sup> Congress has expressed this viewpoint in the E-Government Act of 2002, Public Law No. 107-357. §205(c)(3) of the Act requires the Supreme Court to prescribe rules “to protect privacy and security concerns relating to electronic filing of documents and the public availability ... of documents filed electronically.” Proposed Fed. R. Civ. P. 5.2 has been promulgated in response to this mandate. It should be noted that the E-Government Act and Proposed Fed. R. Civ. P. 5.2 address only issues of personal privacy and do not extend similar rule-based protections to business confidences or proprietary information. Strong concern for the protection of such information was expressed by many members of this Working Group, and this is reflected throughout these Guidelines. However, we acknowledge that there is a special role for rulemaking in the area of personal privacy, as well-developed common law doctrines of business confidentiality and proprietary information are often unknown or unavailable to private citizens.

In different jurisdictions, two primary positions appear to be emerging with respect to the privacy issues relating to electronic case files. The first is sometimes referred to by the shorthand expression, “public is public.” This position assumes that the medium in which case files are stored does not affect the presumption that there is a right of public access. By this analysis, current mechanisms for protecting privacy and confidentiality – primarily through protective orders and motions to seal – are adequate even in the new electronic environment. Some have also suggested that the focus for access policies should be on determining whether information should be deemed “public” in any format – electronic or paper – rather than on limiting access to electronic case files.

Advocates of this position suggest that litigants do not have the same expectation of privacy in court records that may apply to other information divulged to the government. The judicial process depends on the disclosure, voluntarily or involuntarily, of all relevant facts, to allow a judge or jury to make informed decisions. In bankruptcy cases, for example, a debtor must disclose a Social Security number or taxpayer identification number and detailed financial information that the bankruptcy trustee needs to administer the case and that creditors need to fully assert their rights. Similarly, in many types of civil cases – for example, those involving personal injuries, criminal allegations, or the right to certain public benefits – case files often must contain sensitive personal information. To a certain extent, then, advocates of this position expect private litigants to abandon a measure of their personal privacy at the courthouse door.

A second position on the privacy issue focuses on the relative obscurity of paper as compared to electronic files. Advocates of this position observe that unrestricted Internet access undoubtedly would compromise privacy and, in some situations, could increase the risk of personal harm to litigants or others whose private information appears in case files. In other cases, proprietary, financial, marketing, or trade secret information must be disclosed. In these cases, advocates for this position urge that the reasons and justification for public access be balanced against the legitimate needs of corporate litigants to have access to the courts for resolution of disputes without having to forfeit valuable proprietary information to the public and competitors.

The combination of electronic filing and remote access magnifies the potentially dire consequences of mistaken exposure of sensitive information. The accidental disclosure of such information cannot be reversed – mistaken dissemination on the Internet is fundamentally different from an inadvertent disclosure on paper in a courthouse. This reality increases the burden on attorneys and courts to carefully guard against such mistakes. It also has been noted that case files contain information on third parties who often are not able – or not aware how – to protect their personal privacy or to protect valuable proprietary information by seeking to seal or otherwise restrict access to sensitive information filed in litigation.

Advocates of the second position acknowledge that it is difficult to predict how often court files may be used for “improper” purposes in the new electronic environment. They suggest that the key to developing electronic access policies is not the ability to predict the frequency of abuse, but rather the assumption that even a few incidents could cause great personal or competitive harm. Advocates of this position also note that the judicial branch, like other branches of government, has an obligation to protect personal and proprietary information entrusted to it.

They argue that there is a “public interest in privacy” because of the compulsory nature of information disclosure in the context of litigation. That is to say, confidential information often is disclosed in litigation not by choice but by compulsion. On this view, the courts should explicitly recognize, in rules and policies on public access to court records, that although there may not be an expectation of privacy in case file information when there is no protective order, statutory right, or established court procedure that provides such protection, there has been an expectation of practical

obscurity that will be eroded through the development of electronic case files. Appropriate limits on electronic access to certain file information may allow the courts to balance these interests in the context of the new electronic environment.

### **Emerging Themes in the Development of Electronic Access Policy and Procedures**

In efforts to analyze the issues of privacy, confidentiality, and access to electronic court records, the courts are engaging in a debate that in many ways mirrors the broader societal debate over privacy in the Internet era. In the policy development process, courts are addressing two related questions. First, what is the appropriate role of the courts in collecting and maintaining public records? Second, have those courts that allow Internet access to case files changed their role from being passive “custodians” of court records to being active “publishers” of information? These key questions have motivated courts at both the federal and state levels to begin the development of new access policies in the context of electronic case files.

In addition, court policy-making committees also have begun to ask whether the reliance on a case-by-case approach to access issues should be reexamined in the context of Internet publication of court records: Is it prudent to rely on litigants as the primary means of protecting privacy and confidentiality in the context of case files? Judges, as a general matter, do not raise privacy or confidentiality issues on their own. Instead, privacy and confidentiality issues that might be asserted in the course of litigation historically have been addressed on a case-by-case basis, so that if a litigant does not challenge the entry of sensitive information into the record, it will be entered without further inquiry.

Many courts appear to be searching for an alternative to the case-by-case approach, crafting restrictions on remote public access to preserve an element of the practical obscurity of paper files while allowing the public to take advantage of rapid advances in technology to provide easier and cheaper ways to monitor the courts and particular cases. This search for an alternative has led several courts to propose or implement new “categorical” restrictions on access, in effect reversing the common law presumption of access either by presumptively sealing certain types of cases or categories of information or by maintaining open access at the courthouse but restricting remote access on the Internet. In the federal courts, for example, the Judicial Conference of the United States has developed a privacy policy that allows unlimited public access to Social Security case files only at the courthouse, but prohibits remote public access over the Internet. Minnesota has proposed a twist on “courthouse only” access, providing remote public access only to documents and information created or maintained by the courts themselves. Under the Minnesota proposal, documents created by litigants would only be accessible from the courthouse. Other new state court rules on public access – such as those from California, Maryland, and Vermont – carve out limited categories of cases or information for presumptive sealing, adding new categories to existing statutory sealing requirements.

Finally, courts are increasingly focused on “logistics” issues such as data security, the proliferation of electronic documents, and the mechanics of implementing new sealing requirements or access restrictions in the context of electronic case files.

### **Sedona Working Group Assessment of the Main Approaches to Public Access Rules and the Most Common Rule and Policy Features**

Recently-developed rules and policies on public access to court records appear to follow four basic policy approaches. In addition, a preliminary review reveals several issues that each court system seeks to address in developing these rules and policies. Although the Working Group has concluded that it is too soon to identify “best practices” in this area, it is helpful to assess how the new public access rules are consistent – or not – with the principles articulated in these guidelines.



## The Four Basic Policy Approaches

### 1. Open electronic access, with minimal limits.

Some court systems have developed rules or implemented new public access policies affirming that the existing public access system for paper records will continue to apply to electronic court records. This group, which includes the New York state courts and the U.S. federal courts, draws few distinctions between access to electronic and paper files. As a general matter, these court systems only restrict public access to certain personal identifying information in court records that may be used to facilitate the crime of identity theft, but they allow electronic public access to almost all unsealed court records.

Under this approach, the litigant has the obligation to protect private or confidential information by seeking to seal court records or by other self-help mechanisms such as redaction or refraining from filing such information unless absolutely necessary.

This approach, while consistent with the law as it developed before the advent of electronic case filing and potential Internet access to court records, does not take into account the shifting role of courts from “custodians” of records to “publishers,” and engages in no examination of its consequences or desirability. With only minimal restrictions on electronic public access, this shift in the role of the courts may not provide adequate protections for private or proprietary information in court records, and may not be necessary to fulfill the courts’ role in providing appropriate public access to court records.

### 2. Generally open electronic access, coupled with more significant limits on remote electronic public access.

A second group has adopted a middle ground that generally allows remote electronic public access, but at the same time places significant limits on the types of cases – or categories of information – that courts may make available electronically. These courts recognize that there are practical and policy reasons to be cautious about electronic public access, especially in the short-term future during the period of transition from paper to electronic court files. The California and Indiana access rules provide examples of this approach. As with approach #1 above, the obligation to protect personal privacy, confidentiality, or proprietary information will be largely left to individual litigants with little or no independent assistance of the court.

The Working Group views this approach as more balanced than approach #1, but suggests that it may still fall short of providing appropriate protections for private and confidential information in court records. One expects the court to take a more active role, recognizing that public access to court records should be restricted where forfeiture of privacy or proprietary rights would likely result from disclosure.

### 3. Electronic access only to documents produced by the courts.

A third group of courts permits remote access to documents created by the court, such as dockets and court orders, but does not permit remote access to documents created by the parties. This approach is adopted in the Minnesota and Vermont access rules.

The Working Group notes that while this approach appears to provide significant protections for private and confidential information submitted by the litigants as part of the court record, this approach does not appear to allow the public to conduct sufficient oversight of the courts. In many cases, it is important to review pleadings and other litigant-filed elements of the case file in order to effectively evaluate a court’s decision and orders.

### 4. Systematic reevaluation of the content of the public case file, combined with limited access to electronic files.

A fourth broad policy approach is to systematically review the elements of the public case file with the policy goal of better accommodating personal privacy interests in the context of electronic court records. These courts seek to limit the filing of extraneous personal or confidential information in public court files, a strategy referred to as “minimization.” Where such information must be filed, these courts would provide expanded protections before moving forward with electronic public access systems. The Florida court system has issued a report outlining this approach.

The Working Group views this as a promising approach because it focuses on limiting the filing of information that arguably should not be in the public case file, and on sealing or otherwise limiting public access to information that is truly private or confidential, yet also necessary for the adjudication of the dispute.

### **List of Common Features of Recently-Developed Court Rules and Policies on Public Access to Court Records**

The Working Group concludes that, as a general matter, the common features of public access rule outlined below are consistent with the Principles in Chapter 2 on “Pleadings, Court Orders, Substantive Motions, and Dockets,” and Chapter 3 on “Trials.” As noted in those chapters, it is clear that there are some categories of cases that invariably involve information that should be subject to limited public disclosure. State law often mandates that such categories of cases be closed to the public. In addition, many cases involve personal information, the public disclosure of which may violate recognized privacy rights or expose litigants to identity theft or other abuse. In those cases, exclusion of such information from pleadings – or the redaction of such information as mandated in several court rules – will be necessary to protect litigants’ privacy interests while minimally intruding upon the public’s qualified right of access to judicial records.

1. A statement of the overall purpose for the rule or policy.
2. Definitions of key terms used in the rule.
3. A procedure to inform litigants, attorneys, and the public that (a) every document in a court case file will be available to anyone upon request, unless sealed or otherwise protected; (b) case files may be posted on the Internet; and (c) the court does not monitor or limit how case files may be used for purposes unrelated to the legal system.
4. A statement affirming the court’s inherent authority to protect the interests of litigants and third parties who may be affected by public disclosure of personal, confidential, or proprietary information..
5. A list of the types of court records that are presumptively excluded (sealed) from public access by statute or court rule.
6. A statement affirming that the public right to access court records and the court’s authority to protect confidential information should not, as a general matter, vary based on the format in which the record is kept (e.g., in paper versus electronic format), or based on the place where the record is to be accessed (i.e., at the courthouse or by remote access).
7. As an exception to feature 6 above, a list of the types of court records that – although not sealed – will not be available by remote electronic public access.
8. A list of the types of information that either: a) must not be filed in an open court record, or b) if filed, must be redacted or truncated to protect personal privacy interests. These provisions mainly apply to personal identifiers such as the SSN, account numbers, and home addresses of parties.

9. Procedure for a court to collect and maintain sensitive data elements (such as SSN) on special forms (paper or electronic) that will be presumptively unavailable for public access. Such procedures generally build on technology to segregate sensitive information so that public access can be restricted in appropriate situations.
10. Procedure to petition for access to records that have been sealed or otherwise restricted from public access, and a statement of the elements required to overcome the presumption of non-disclosure.
11. Procedure to seal or otherwise restrict public access to records, and a statement of the burden that must be met to overcome the presumption of disclosure.
12. An affirmation that a rule on public access to court records does not alter the Court's obligation to decide, on a case-by-case basis, motions to seal or otherwise restrict public access to court records.
13. Guidance to the courts concerning data elements that are contained in electronic docketing systems that must (or must not) be routinely made available for public access.
14. Guidance for attorneys and/or litigants concerning: (a) the extent to which public case files will be made available electronically; and (b) the need to exercise caution before filing documents and information that contain sensitive private information, which is generally defined elsewhere in the rule.
15. An explanation of the limits, if any, on the availability of "bulk" and/or "compiled" data from public court records. Some rules specify that such data will only be made available to certain entities, for certain defined purposes, and pursuant to agreements to refrain from certain uses of the records obtained.
16. A statement concerning the fees that a court may charge for public access to court records.

## **Conclusion**

Courts have begun to address privacy and confidentiality issues that arise as court files are made accessible on the Internet. The federal courts and a growing number of state court systems have developed policies or court rules to balance the competing interests of public access and personal privacy. These policies and rules recognize that case files may contain sensitive personal, confidential or proprietary information that may require special protection in the context of Internet access. Through changes in rules, court policies – and likely also in case law – it is clear that the law in this area will continue to develop to respond to the fundamentally changed context of public access to court records in the Internet era.

It is also clear that the era of "practical obscurity" has past. Litigants and attorneys must be aware of the possible consequences of filing any private or confidential information.

## CHAPTER 5 Selected Bibliography

### *Court Rules and Policy Statements*

Federal Judicial Conference Privacy Policy, *available at* <http://www.privacy.uscourts.gov/b4amend.htm>

United States District Court, District of New Jersey, L. Civ. R. 5.3

California Special Rules for Trial Courts (Rules 2070 to 2077), *available at* <http://www.courtinfo.ca.gov/rules/titlefive/titlefive.pdf>

Indiana Administrative Rules (Rule 9), *available at* <http://www.in.gov/judiciary/orders/rule/amendments/2004/0204-admin-rule9.pdf>

Supreme Court of Florida, In re: Implementation of Report and Recommendations of the Committee on Privacy and Court Records, Administrative Order No. AOSC06-20, June 30, 2006, *available at* <http://www.floridasupremecourt.org/clerk/adminorders/2006/sc06-20.pdf>

Supreme Court of Florida, In Re: Interim Policy in Electronic Release of Court Records, Administrative Order No. AOSC06-21, June 30, 2006, *available at* <http://www.floridasupremecourt.org/clerk/adminorders/2006/sc06-21.pdf>

Maryland Rules of Procedure (Ch. 1000 – Access to Court Records), *available at* <http://courts.state.md.us/access/ro-accessstocrecords.pdf>

Minnesota Recommendations of the Minnesota Supreme Court Advisory Committee on Rules of Public Access to Records of the Judicial Branch, *available at* [http://www.courts.state.mn.us/cio/public\\_notices/accessreport.htm](http://www.courts.state.mn.us/cio/public_notices/accessreport.htm)

New York Commission on Public Access to Court Records, Report to the Chief Judge of the State of New York, *available at* <http://www.courts.state.ny.us/ip/publicaccess/index.shtml>

Vermont Rules for Public Access to Court Records (and amendments), *available at:*  
[http://www.vermontjudiciary.org/rules/proposed/rules\\_pa.htm](http://www.vermontjudiciary.org/rules/proposed/rules_pa.htm)  
[http://www.vermontjudiciary.org/rules/proposed/rules\\_pamay1.htm](http://www.vermontjudiciary.org/rules/proposed/rules_pamay1.htm)  
<http://www.vermontjudiciary.org/rules/proposed/paAmend6b.htm>

Rules Governing Dissemination of Electronic Case Records, *available at* <http://www.vermontjudiciary.org/rules/proposed/ruleselectronic.htm>

### *Cases*

United States Supreme Court

*U.S. Dept. of Justice v. Reporters Committee for Freedom of the Press*, 489 U.S. 749 (1989) (interpreting the FOIA, recognizing a privacy interest in information that is publicly available through other means, but is “practically obscure”).

***United States District Courts***

*Holland v. GMAC Mortgage Corp.*, 2004 WL 1534179 (D. Kan. Jun 30, 2004) (noting that filing documents under seal in court which employs electronic filing will be burdensome because sealed filing must be in paper format).

*Eldagbar v. City of N.Y. Dep 't of Citywide Admin. Servs.*, 2004 U.S. Dist. LEXIS 3503 (S.D.N.Y. 2004) (denying motion for protective order, observing there is no protectable privacy interest in personal identifiers that are already available on Internet).

***Websites***

National Center for State Courts

<http://www.courtaccess.org/>

Reporters' Committee for Freedom of the Press

<http://www.rcfp.org/courtaccess/viewstates.php>

*A full bibliography, updated periodically, may be found on The Sedona Conference web site at [www.thesedonaconference.org](http://www.thesedonaconference.org)*

# THE SEDONA CONFERENCE® BEST PRACTICES COMMENTARY ON THE USE OF SEARCH AND INFORMATION RETRIEVAL METHODS IN E-DISCOVERY

---

*A Project of The Sedona Conference® Working Group on Best Practices for Document Retention and Production (WG1), Search & Retrieval Sciences Special Project Team\**

*August 2007 Public Comment Version*

Editor-in-Chief:  
Jason R. Baron

Executive Editors:  
Richard G. Braman  
Kenneth J. Withers

Senior Editors:  
Thomas Y. Allman  
M. James Daley  
George L. Paul

---

\* With valuable input from many other WG1 members and the RFP+ Vendor Panel. This document is for educational purposes only and is not a substitute for legal advice. The opinions expressed herein are consensus views of the editors and authors, and do not necessarily represent the views of any individual participants or authors or any of the organizations to which they belong or clients they represent, nor do they necessarily represent official views of The Sedona Conference®.

*Table of Contents*

<i>Preface and Acknowledgements</i> .....	191
<i>Overview</i> .....	192
<i>Executive Summary</i> .....	193
I. Introduction .....	196
II. The Search and Information Retrieval Problem Confronting Lawyers .....	197
III. Lawyers' Present-day Use of Search and Retrieval Methodologies .....	199
IV. Some Key Terms, Concepts and History in Information Retrieval Technology .....	204
V. Boolean and Beyond: A World of Search Methods, Tools and Techniques .....	207
VI. Practical Guidance in Evaluating the Use of Automated Search And Retrieval Methods .....	208
VII. Future Directions in Search and Retrieval Science .....	212
Appendix: Types of Search Methods .....	217

---

*Preface and Acknowledgements*

---

Welcome to another major publication in The Sedona Conference® Working Group Series (the “WGS”®), *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*. This effort is an outgrowth of our Working Group on Electronic Document Retention and Production (WG1) and represents the work of its Search and Retrieval Sciences Special Project Team, consisting of a diverse group of lawyers and representatives of firms providing consulting and legal services to the legal tech community.

The mission of the Search and Retrieval Sciences Special Project Team has been to explore the nature of the search and retrieval process in the context of civil litigation and regulatory compliance in the digital age. The goal of this Best Practices Commentary is to provide the bench and bar with an educational guide to an area of e-discovery law that we believe will only become more important over time, given the need to accurately and efficiently search for relevant evidence contained within the exponentially increasing volumes of electronically stored information (ESI) that are stored and made subject to litigation, investigations, and regulatory activities. We also understand that the subject of what constitutes best practices in this area will necessarily be subject to change, given the accelerating pace of technological developments that the law is struggling to keep up with. We hope that our efforts will assist the legal profession in this area, and we welcome all feedback at [tsc@sedona.net](mailto:tsc@sedona.net).

This Commentary was originally conceived at the Fourth Annual Meeting of WG1 in Vancouver, B.C., in the fall of 2005. Through the efforts of many individual contributors and editors, several successive drafts were prepared for comment by the full WG1 membership in successive mid-year and annual meetings. I especially want to acknowledge the contributions to the overall success of this project made by Jason R. Baron, who took the lead role in editing the Commentary, along with all of the special contributions of his fellow Co-chairs of the Search and Retrieval Sciences Team, M. James Daley and Ariana J. Tadler. I also wish to acknowledge the invaluable editorial assistance provided on one or more successive drafts by senior contributing editors Thomas Y. Allman, M. James Daley, and George L. Paul, as well as the drafting contributions provided along the way by Macyl Burke, Christopher Cotton, Matthew Cohen, Conor Crowley, Sherry Harris, William Herr, Stephanie Mendelsohn, Dan Regard, Herbert Roitblat, Sonya Sigler, and Stephen Whetstone. Lastly, I wish to acknowledge that many other individuals in WG1, including on the Search and Retrieval Sciences Special Project Team and the RFP+ Vendor Panel, spent time in collaborating on earlier proposals for material to be included in the Commentary. On behalf of Richard Braman, Executive Director of The Sedona Conference®, I wish to thank everyone involved in devoting their time and attention during the drafting and editing process.

*Kenneth J. Withers*

Director, Judicial Education and Content  
The Sedona Conference®  
June 2007



***Traditional Approaches To Searching For Relevant Evidence  
Are No Longer Practical Or Financially Feasible***

Discovery of the relevant information gathered about a topic in dispute is at the core of the litigation process.<sup>1</sup> 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 poses unique challenges. Just a few years ago, a party seeking to review information for production to the other side in a “large” document review case might have been concerned with hundreds of “banker’s” boxes of documents.

Today, that same amount of data might be found on a single computer hard drive.<sup>2</sup> 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 typical gigabyte of storage cost about \$20,000; today it costs less than \$1 dollar. As a result, more individuals and companies are generating, receiving and storing more data, which means more information must be gathered, considered, reviewed and produced in litigation. But, with billable rates for junior associates at many law firms now starting at over \$200 per hour, the cost to review just one gigabyte of data can easily exceed \$30,000.<sup>3</sup> These economic realities – *i.e.*, the huge cost differential between the \$1 to store a gigabyte of data and the \$30,000 to review it – act as a driver in changing the traditional attitudes and approaches of lawyers, clients, courts and litigation support providers about how to search for relevant evidence during discovery and investigations. Escalating data volumes into the billions of ESI objects, review costs, and shrinking discovery timetables, all add up to equaling the need for profound change.

As discussed below in this Commentary, just as technology has given rise to these new litigation challenges, technology can help solve them, too. The emergence of new discovery strategies, best practices and processes, as well as new search and retrieval technologies, are transforming the way lawyers litigate and, collectively, offer real promise that huge volumes of information can be reviewed faster, more accurately, and more affordably than ever before. The good news is that search and retrieval systems are improving and expanding, buoyed by a huge economic wave of activity aimed at improving the “search” experience for users generally.<sup>4</sup> For example, advanced forms of search techniques, including various forms of fuzzy logic, text mining and machine learning all automatically organize electronically stored information in new ways not achieved by past more familiar methods, including the simple use of “keywords” as the only automated aid to conducting manual searches. Although we are at the dawn of a new era, these new techniques hold the potential to increase both accuracy and efficiency. Through statistical sampling and validation techniques we can then confirm 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 serve as a guide to enable both the bench and the bar to become more familiar with the new challenges presented by needing to search and retrieve electronically stored information. The Commentary seeks to identify ways to address those challenges, and select the best solution to maximize the just, speedy, and inexpensive determination of every action, consistent with Federal Rule of Civil Procedure 1.

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 70,000-80,000 of text pages, or 35 to 40 banker’s boxes of documents (at 2,000 pages per box). Thus, a 100-gigabyte storage device (*e.g.*, a personal computer hard drive), theoretically, could hold as much as the equivalent of 3,500 to 4,000 banker’s boxes of documents. By contrast, in 1990, a typical personal computer held just 200 megabytes of data - 1/500 the capacity of a typical hard drive today. Even if only 10% of a computer’s available capacity today contains useful or “useable” information (as distinguished from application programs, operating systems, utilities, etc.), attorneys still would need to consider and potentially review 700,000 to 800,000 pages per each device.

3 See Commentary, *infra*, n.13.

4 One indication of the amount of ongoing effort and investment generally to improve search and retrieval capabilities is evidenced by the research and development spending of internet giants Google, Yahoo!, and eBay. According to published reports, Google spent \$ 1.23 billion, Yahoo! spent \$883 million, and eBay spent \$495 million on core research and development activities in 2006. See Robert Hertzberg, “LT’s Top 84 R&D Spenders,” *Baseline* (April 17, 2007), [www.baselinemag.com/article2/0,1540,2114821,00.asp](http://www.baselinemag.com/article2/0,1540,2114821,00.asp).

---

*Executive Summary*

---

Discovery has changed. In just a few years, the review process needed to identify and produce information has evolved from one largely involving the manual review of paper documents to one involving vastly greater volumes of electronically stored information.

A perfect review of the resulting volume of information is not possible. Nor is it economic. The governing legal principles and best practices do not require perfection in making disclosures or in responding to discovery requests.

The Sedona Conference<sup>®</sup> 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 evolutionary 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 – continues to advance, a new understanding will evolve about what is “reasonable” under the particular circumstances of those technologies. Thus, the challenges addressed by this Commentary go beyond litigation and encompass all aspects of the search and retrieval of information from large volumes of data.

### **The Revolution in Discovery**

Just a few years ago all information was stored on physical records such as paper. There was typically only one original document, and the number of duplicative copies and their location was generally limited. Administrative assistants, file clerks, records managers and archivists developed expertise in managing the storage, generally pursuant to pre-existing file systems. It was reasonable, and indeed relatively easy in all but the exceptional case, for the legal profession to gather and then manually review all the individual items collected as part of the discovery process prior to their production.

But with the digital revolution there has also been a paradigm shift in the review process which is feasible. The shift of information storage to a digital realm has, for a variety of reasons, caused an explosion in the amount of information that resides in any enterprise-profoundly affecting litigation. This massive amount of electronically stored information is distributed broadly among different storage devices, from large mainframe computers, to tiny machines capable of storing information equivalent to several warehouses of documents each, all of which are or can be integrated into other systems. These systems are complex, interdependent, and evolve spontaneously, like ecosystems. It is often impossible to find one person, or even one discrete group of people, who completely understand the workings of this new form of “information ecosystem.”

Finally, added to the search and retrieval challenge is the fact that a large percentage of the records being searched are expressed in *human language*, not just numbers. Human language is an inherently elastic, ambiguous “living” tool of enormous power. Its elasticity allows for private codes and vocabularies to exist in different subcultures in any enterprise, thus making the identification of the “words” to be searched much more challenging.

### **Essential Conclusions of this Commentary**

This Sedona Conference<sup>®</sup> “Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery” strives to set forth state-of-the-art knowledge about

meeting the challenge of searching enormous databases for relevant information, and then retrieving that information with a minimum of wasted effort.

By way of summary, we set forth our conclusions about the Problems and their Solutions, and summarize our Practical Advice which the balance of the paper articulates.

### Problems

- Exponential growth in informational records is a critical challenge to the justice system.
- Electronically stored information contains human language, which challenges computer search tools. These challenges lie in the ambiguity inherent in human language and tendency of people within organizations or networks to invent their own words or communicate in code.
- The comparative efficacy of the results of manual review versus the results of alternative forms of automated methods of review remains very much an open matter of debate. Moreover, simple keyword searching, while itself a valuable tool, has certain known deficiencies.

### Solutions

- Much that is useful in selecting information for production in discovery can be learned from other disciplines, including: information retrieval science; the study of linguistics; and implementation of effective management processes, to name just a few.
- Alternative search tools are available to supplement simple keyword searching and Boolean search techniques. These include using fuzzy logic to capture variations on words; using conceptual searching, which makes use of taxonomies and ontologies assembled by linguists; and using other machine learning and text mining tools that employ mathematical probabilities.
- It may be useful and appropriate to seek agreement on ways to measure and evaluate the effectiveness of the search and retrieval process. The metrics currently used in information science, such as “precision” and “recall,” as well as more involved concepts are worth studying.

### Practical Advice

*Practice Point 1. In many settings involving electronically stored information, reliance 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.*

*Practice Point 2. Success in using 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.*

*Practice Point 4. Parties should perform due diligence in choosing a particular information retrieval product or service from a vendor.*

*Practice Point 5. The use of search and information retrieval tools does not guarantee that all responsive documents will be identified in large data collections, due to characteristics of human language. Moreover, differing search methods may produce differing results, subject to a measure of statistical variation inherent in the science of information retrieval.*

- Practice Point 6. Parties should make a good faith attempt to collaborate on the use of particular search and information retrieval methods, tools and protocols (including as to keywords, concepts, and other types of search parameters).*
- Practice Point 7. Parties should expect that their choice of search methodology will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and trials).*
- Practice Point 8. Parties and the courts should be alert to new and evolving search and information retrieval methods.*

### **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 growing consensus that the application of linguistic and mathematic-based content analysis, embodied in new forms of search and retrieval technologies, tools, techniques and process in support of the review function can effectively reduce litigation cost, time, and error rates.

### **Recommendations**

- 1. The legal community should 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, for use in assisting lawyers in evaluating 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 fora focused on issues involving information retrieval. 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.

## I. INTRODUCTION

The exponential growth in the volume of electronically stored information or “ESI” found in modern enterprises poses a substantial challenge to the justice system. Today, even routine discovery requests can require searches of the storage devices found on mainframes, servers, networked workstations, desktops and laptops, home computers, removable media (such as CDs, DVDs and USB flash drives), and handheld devices (such as PDAs, cell phones and iPods). Complicating things, 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, de-fragmented, morphed and multiplied. Discovery requests for e-mail, as one common example of ESI, often require searching and retrieving information from thousands to millions or even tens of millions of individual messages, with attachments in various file formats.

The volume and complexity of this electronically stored information highlights several issues: First, whether automated search and information retrieval methods are reliable and accurate? Second, whether the legal profession has developed the skills, know-how and processes to use such automated search and retrieval methods intelligently, when applied to huge data sets, in ways that are defensible under the rules governing discovery? Yet another issue is what impact, if any, the changes to the Federal Rules governing e-discovery will have on the search and retrieval process?

The Sedona Principles, Second Edition (2007) issued by The Sedona Conference<sup>®</sup> have endorsed several highly pragmatic and relevant consensus best practices relevant to this discussion.<sup>5</sup>

First, Principle 6 provides that responding parties are in the best position “to evaluate the procedures, methodologies, and technologies appropriate or 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 goes on to state 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.”<sup>6</sup> 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 agreement should take account of the iterative nature of the discovery process and allow for refinement as the parties’ understanding of the relevant issues develops.”<sup>7</sup>

Third, the Sedona Conference<sup>®</sup> has recognized that “there are now hundreds of companies offering electronic discovery services.”<sup>8</sup> This is also true of search and 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 substantial 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 – about “state of the art” search and retrieval tools, techniques, and methodologies, and

---

5 *The Sedona Principles, Second Edition: Best Practices Recommendations & Principles for Addressing Electronic Document Production (The Sedona Conference® Working Group Series, 2007)* (“*The Sedona Principles, Second Edition, 2007*”), available at [www.thosedonacorporation.com](http://www.thosedonacorporation.com).

6 *Id.*, Comment 11.a.

7 *Id.*

8 *The Sedona Conference® Best Practices for the Selection of Electronic Discovery Vendors: Navigating the Vendor Proposal Process* (2007), available at [http://www.thosedonacorporation.com/content/miscFiles/RFP\\_Paper.pdf](http://www.thosedonacorporation.com/content/miscFiles/RFP_Paper.pdf).

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.

The legal community is familiar with keyword and natural language searches on Westlaw<sup>®</sup> and Lexis<sup>®</sup> 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. However, the use of keyword, Boolean, and other search and retrieval tools to narrow information to be reviewed for production in discovery is relatively recent.<sup>9</sup> Moreover, to date, the relative efficacy of competing search and retrieval tools used to accomplish production review simply have not been measured. The field is wide open for the development of search and information retrieval best practices that take into account various alternative search and retrieval methods. These methods extend from improvements in basic keyword searching, to more sophisticated systems that use 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 describes the world of search tools, techniques and methodologies that are currently commercially available. It also includes a “practice pointers” guide 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 updated Sedona Glossary.

## II. THE SEARCH AND INFORMATION RETRIEVAL PROBLEM CONFRONTING LAWYERS

The discovery process of today is drowning in potential sources of information. The exponential increase in volume, especially since the mid-1990s, is principally due to the impact of the PC revolution, the widespread use of email and the growth of 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 a Crushing Volume 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 30 years, there has been a fast-paced and widespread shift from civilization’s original 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 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

---

<sup>9</sup> There may be a role for 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” phase, 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 a production phase review. 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 is truly 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 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 RICHMOND J. LAW & TECHNOLOGY 11 (2007), at Paragraphs 53, 60, <http://law.richmond.edu/jolt/v13i3/article11.pdf> (discussing the use of concept searching in regard to preservation); *The Sedona Principles, Second Edition*, 2007, Comment 11.a (“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.”).

personal computer, the spread of various types of networks linking together both computers and other networks, the rise of e-mail 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.<sup>10</sup>

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 happening on a 24/7 basis. For example, studies reflect that the average U.S. worker sends and receives 100 e-mails per day. The size and nature of the attachments to these emails is also growing, with increased integration of image, audio and video files. Most recently, there has been a similar explosion in the use of instant messaging throughout business enterprises. In many organizations, the average worker maintains several gigabytes of stored data.<sup>11</sup> At the same time, the costs of storage have plummeted from \$20,000 per gigabyte in 1990 to less than \$ 1 per gigabyte today.<sup>12</sup> 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 ESI world.

Companies have continued to aggressively leverage technology to increase productivity. No one really controls how, where, how many times, and in how many forms information is stored. For example, the same Word documents can be found on e-mail attachments, local hard drives, network drives, document management systems, websites, and on all manner of removable media, such as USB flash drives, CDs, DVDs, and so on.

### ***Discovery During the Recent Past: Manageable Amounts of Physically Stored Information***

Historically, outside counsel played a key role in the discovery process, and the process worked simply. Litigants, assisted by their counsel, identified and collected information that was relevant to pending or foreseeable litigation. Counsel reviewed the information and produced any information that was relevant 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 had been created in or was stored in printed, physical form, and in reasonable volumes so that it required only “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 has become neither workable nor 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 electronically stored information in the time permitted for discovery.<sup>13</sup> Moreover, efforts to reduce time and cost by use of “claw back”<sup>14</sup> provisions are problematic because of the risk of disclosure of sensitive proprietary and privileged information, as well as the risk of privilege waiver that can be imposed by substantive law, irrespective of new changes in procedural rules.

Accordingly, the conventional discovery review process is poorly adapted to much of today’s litigation.<sup>15</sup> Lawyers of all stripes therefore have a vital interest in utilizing automated search and

10 See George L. Paul and Jason R. Baron, “Information Inflation: Can the Legal System Adapt?,” 13 RICHMOND J. LAW & TECHNOLOGY 10 (2007), at Paragraph 1, n.2, <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, <http://www.sims.berkeley.edu/how-much-info-2003>.

11 As noted *supra*, n.2, one gigabyte is equivalent in volume to between 70,000 and 80,000 pages of material. At 2000 pages per box, one gigabyte is therefore equivalent to 35 to 40 boxes of documents.

12 Michelle Kessler, “Days of officially drowning in data almost upon us,” *USA Today*, Mar. 5, 2007, available at [www.usatoday.com/tech/news/2007-03-05-data\\_N.htm](http://www.usatoday.com/tech/news/2007-03-05-data_N.htm).

13 Compare \$1 to store a gigabyte of data with \$32,000 to review it (*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).

14 “Clawback” and “quick peek” provisions in case management agreements seek to permit large productions of electronic data little or no review, and without waiver of any claim of privilege, work product, etc. See *The Sedona Principles, Second Edition, 2007*, Comment 10.d. See also amended Fed. R. Civ. P. 26(f)(4), effective December 1, 2006, and accompanying Committee Note.

15 Not all cases are equally heavy in involving electronic discovery and, from time to time, counsel may forgo production of electronically stored information and rely solely on hard copy documents.

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 maximizing the quality of discovery, by means of using automated tools that produce a reliable, reproducible and consistent product.

Ideally, then, judges and litigants should strive to increase their awareness of search and retrieval sciences generally, and of their appropriate application in 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 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 review of massive sets of electronic data sets (which it does not), the relative efficacy of that approach versus utilizing newly developed automated methods of review remains very much open to debate. Moreover, past research demonstrates the gap between lawyers' expectations and the true efficacy of certain types of searches. The Blair and Maron study (discussed below) reflects that human beings are less than 20% to 25% accurate and complete in searching and retrieving information from a heterogeneous set of documents (*i.e.*, in many data types and formats). 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 yet be of benefit in litigation.

### ***The Intelligent Use of Tools***

Although the continued use of manual search and review methods may be indefensible in discovery involving significant amounts of electronically stored information, merely adopting sophisticated automated search tools, alone, will not necessarily lead to successful results. Lawyers must recognize that, just as important as utilizing the automated tools, is tuning the *process* in and by which a legal team uses such tools, including a close involvement of lead counsel. This may require an iterative process which importantly utilizes feedback and learning as tools, and allows for measurement of results. The time and effort spent on the front end designing a sophisticated discovery process that targets the real needs of the client 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 years to using the automated means of searching provided by LexisNexis® and Westlaw® databases. More recently, lawyers have begun to use Google® and other Web-based search engines to hunt down information relevant to their practice. Additionally, law firms and corporate legal departments use search methods for administrative matters, such as searching data on available personnel, to support billing functions, to manage conflicts of interest, and for purposes of 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 their opposing counsels' pleadings,



motions, and pretrial filings in similar litigation, as well as showing how a judge has ruled in similar issues even if unreported in legal reporting services. Lawyers also use a variety of search methods with online and CD-ROM databases to dig up facts on opposing parties, witnesses, and even jury pools. At later stages of litigation, lawyers use various litigation management 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.

### “De-duplication” in the Processing of ESI

With the exponential increase in the amount 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 “de-duplication” 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. Such de-duplication software reduces the time attorneys must spend reviewing a large document set and helps to ensure consistent classification of documents for responsiveness or privilege.<sup>16</sup> Increasingly, “near de-duplication” tools also are being 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.<sup>17</sup>

### The Use of “Keywords”

By far the most commonly used search methodology today is 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 represented a significant advance in using automated technologies, as increasingly recognized by the courts. As one United States Magistrate Judge stated, “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.”<sup>18</sup>

Courts have not only accepted, but in some cases have ordered, the use of keyword searching to define discovery parameters and resolve discovery disputes. One court has also suggested that a party might satisfy its duty to preserve documents in anticipation of litigation by conducting system-wide keyword searching and preserving a copy of each “hit.”<sup>19</sup>

Because of the costs and burdens (if not impossibility) of reviewing increasingly vast volumes of electronic data, it makes sense for producing parties to negotiate with requesting parties in advance to define the parameters of discoverable information. For example, parties could agree on

16 “De-duplication” services work to tag identical documents as duplicates by means of a “binary hash function” (which simply is 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 in fact perfectly alike). De-duplication 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 de-duplication process); *Medtronic Sofamor Danek Inc. v. Michelson*, 229 F.R.D. 550, 561 (W.D.Tenn., 2003) (same).

17 “Near de-duplication” involves files that “are not hash value duplicates but are materially similar.” See <http://www.law.com/jsp/legaltechnology/roadmapArticle.jsp?id=1158014995345&chubpage=Processing>.

18 *In re Lorazepam & Clonazepam*, 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 6188563 (N.D. Cal. Mar. 13, 2006) (“e-mail could likely be screened efficiently through the use of electronic search terms that the parties agree upon”); *Windy City Innovations, LLC v. American 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 e-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 Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550 (W.D. Tenn. 2003) (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).

19 *Zubulake v. UBS Warburg, LLC*, 229 F.R.D. 422 (S.D.N.Y. 2004); cf. *Caché 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). 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 e-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*).

conducting a search of only files maintained by relevant or key witnesses, and/or for certain date ranges. They often can also agree to a set of key words relevant to the case. Both sides can often 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, which comes with the territory of searching through electronic realms.<sup>20</sup>

In *Treppel v. Biovail Corp.*,<sup>21</sup> 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 recent case emphasized the need to confer after plaintiff was successful in obtaining a "mirror image" of data on all of defendant's computers.<sup>22</sup>

### Issues With Keywords

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, keyword searches work well to find all documents that mention a specific individual or date, regardless of context. However, although basic keyword searching techniques have been widely accepted both by courts and parties as sufficient to 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 end up being both over- and under-inclusive in light of the inherent malleability and ambiguity of spoken and written English (as well as all other languages).<sup>23</sup>

Keyword searches identify all documents containing a specified term regardless of context, and so they can possibly capture 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 positive" hits or noise in the data is potentially huge, amounting in some cases to huge numbers of files which must be searched to find responsive documents.<sup>24</sup>

On the other hand, 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

20 See generally Kenneth J. Withers, *Computer-Based Discovery in Federal Court Litigation*, 2000 FEDERAL COURTS L. REV. 2, <http://www.fclr.org/articles/2000fedcslrev2.pdf> (suggesting parties adopt collaborative strategies on search protocols); see also R. Brownstone, *Collaborative Navigation of the Stormy e-discovery Seas*, 10 RICHMOND J. LAW & TECHNOLOGY 53 (2004), <http://law.richmond.edu/joltr/v10i5/article53.pdf> (arguing that parties must agree to search terms and other selection criteria to narrow the scope of manageable data sets); see also *The Sedona Principles, Second Edition, 2007*, Comment 11.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.").

21 233 F.R.D. 363 (S.D.N.Y. 2006).

22 *Balboa Threadworks v. Sucky*, 2006 WL 763668 (D. Kan. Mar. 24, 2006) (court orders parties to meet and confer on the use of a search protocol, including key word searching).

23 Some case law has held that keyword searches were either incomplete or overinclusive, see *Alexander v. FBI*, *supra*, n.18; *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).

24 See, e.g., G. Paul and J. Baron, "Information Inflation," *supra*, n.10, at Paragraph 20 (discussing potential time and cost of searching through 1 billion emails); Craig Ball, "Unlocking Keywords: How you frame your search words will shape your success," 14 No. 1 *Law Technology News* 56 (January 2007) (discussing how to improve keyword search results by use of various techniques, including eliminating "noise words" such as "law" and "legal"). See also Steven C. Bennett, "E-Discovery by Keyword Search," 15 No. 3 *Prac. Litigator* 7 (2004).

“synonymy” or *variation* in the use of language). And of course, if authors of records are inventing words “on the fly,” as they have done through history, and now are doing with increasing frequency in electronic communications, such problems are compounded.<sup>25</sup>

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”). So too, it is well known that even the best of optical character recognition (OCR) scanning processes introduce a certain rate of random error into document texts, potentially transforming would-be keywords into something else. Finally, using keywords alone results in a return set of potentially responsive documents that are not weighted and ranked based upon their potential importance or relevance. In other words, each document is considered to have an equal probability of being responsive upon 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, however, 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 continue to rely on keyword searching. Such methods are limited by their dependence on matching a specific, sometimes arbitrary choice of language to describe the targeted topic of interest.<sup>26</sup> The issue of whether there is room for improvement in the rate of “recall” (as defined in the next section) of relevant documents in a given collection is something lawyers must consider when relying on simple and traditional input of keywords alone.

### *Use of Alternative Search Tools and Methods*

Lawyers are beginning to feel more comfortable using alternative search tools to identify potentially relevant electronically stored information. These more advanced text mining tools include “conceptual search methods” which rely on semantic relations between words, and/or which 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 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 “Alaska.” At least one reported case has referenced the possible use of “concept searching” as an alternative to strict reliance on keyword searching.<sup>27</sup>

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 communications that might be lurking in documents, files, and emails. For example, a linguist would want to know how union organizers or company officials might

25 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., The Macmillan Co., 1953 (“[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”).

26 See Part IV, *infra*; see generally, S.I. Hayakawa, *LANGUAGE IN THOUGHT AND ACTION* (Harcourt 1990) (5th ed.) (such methods are inherently limited by their specific choice of language to describe a specific object or reality).

27 *Disability Rights Council of Greater Washington v. Washington Metropolitan Transit Authority*, 2007 WL 1585452 (D.D.C. June 1, 2007) (citing to G. Paul and J. Baron, *supra*, n.10); see generally M. Mazza, E. Quesada, and A. Sternberg, “In Pursuit of FRCP 1: Creative Approaches To Cutting and Shifting the Costs of Discovery of Electronically Stored Information,” *supra* n.9, at Paragraph 54 (discussing concept searching).

communicate plans, any special code words used in the industry, the relationships of collective bargaining units, company 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 arguably helpful in addressing cultural biases of taxonomies because they do not depend on linguistic analysis, but on mathematical probabilities. They can also help to find communications 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, used in government intelligence, is particularly apt in helping lawyers find information when they don’t 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.

Anecdotal information suggests that a small number of companies and law firms – particularly those that have gained significant experience in e-discovery – are using alternative search methods to either identify responsive documents (reducing expensive attorney review time) or to winnow collections to the key documents for depositions, pretrial pleadings, and trial.

The document databases that can assist lawyers in developing advanced ontologies and mathematical models are not limited to “discovery” documents. Search tools can be used in overall case management to search across pleadings, legal research, discovery responses, expert reports, and attorney work product. For example, in addition to searching discovery documents, a legal team in a labor dispute might want to search the interrogatory responses, pleadings, and depositions for all references to the concept of “strike.” This is a potential growth area for vendors specializing in case management software.

Apart from the authorities listed in this section, there is still little by way of published reports or cases discussing or challenging the use of these various tools. It is only a matter of time, however, before more widespread deployment will lead to the development of a fuller body of case law.

### *Resistance by the Legal Profession*

Some litigators continue to primarily rely upon manual review of information as part of their review process.<sup>28</sup> Principal rationales are: (1) concerns that computers cannot be programmed to replace the human intelligence required to make complex determinations on 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 searching, yet be reluctant to use alternative search techniques. They may not be convinced that the chosen method would withstand a court challenge. They may perceive a risk that problem documents will not be found despite the additional effort; and an opposite risk 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 a tried-and-true lowest common denominator. Finally, litigators lack the time and resources to sort out these highly complex technical issues on a case-by-case basis.<sup>29</sup>

---

28 *But see In re Instinet Group, Inc.*, 2005 WL 3501708 (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.”

29 *See, e.g., R. Friedmann*, <http://prismlegal.com/wordpress/?cat=9> (Feb. 4, 2005) (suggesting that not one solution fits all cases); *see also id.* (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).

### *Challenging the Choice of Search Method*

The challenge to a choice of 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 use 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 reduce challenges – advocated by the proponents of the 2006 Federal Rules Amendments and experienced practitioners – is for a full and transparent discussion among counsel of the search terminology. 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.

Absence 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 data. However, a unilateral choice of a search methodology may be challenged due to lack of a scientific showing that the results are accurate, complete and reliable. Since all automated search tools rely on some level of science, the challenging party may argue that the process used by the responding party is essentially an expert technology which has not been validated by subjecting it to peer review, and unbiased empirical testing or analysis.

The probability of such a challenge is greater if the technology is patented or proprietary to a developer or vendor (*i.e.*, in a so-called “Black Box”). In such circumstances, e-discovery and litigation support vendors that use these technologies may be several degrees of separation from the original developers. A requesting party may demand the responding party to “prove up” the use of such search technology. This could set the stage for a difficult and expensive battle of experts.

As a practical matter, however, those who might object to a particular search and retrieval technology face several challenges. First, the legal system has, for decades blessed the use of keyword search tools and databases for discovery review. Second, even if such a challenge were permitted to proceed, the lack of a formally acknowledged baseline by which to measure the comparative accuracy and reliability of any search method precludes a comparison of the “new” method to traditional methods. And third, if human review or even keyword searching is the benchmark for accuracy and reliability, it arguably should not be difficult to compare the new technology favorably with either 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, we would expect that a body of precedent will develop over time which references, if not critically analyzes, new and alternative search methods in use in particular legal contexts.

## IV. SOME KEY TERMS, CONCEPTS AND HISTORY IN INFORMATION RETRIEVAL TECHNOLOGY

The evaluation of information retrieval (“IR”) systems has, until now, largely been of greatest interest to computer scientists and graduate students in information and library science. Unlike performance benchmarking for computer hardware, there are no agreed-upon objective criteria for evaluating the performance of information retrieval systems. That is, for IR systems, the notion of effectiveness is subjective. Human judgment is ultimately the criteria 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. Additionally, the subject matter and information type impact a user's information retrieval 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 information retrieval 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.”<sup>30</sup> 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.”<sup>31</sup> Another way to think about it is: out of the total number of relevant documents in the document collection, how many were retrieved correctly?

*Precision* is defined as “an information retrieval performance measure that quantifies the fraction of retrieved documents which are known to be relevant.”<sup>32</sup> 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, 1000 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 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 usually a trade off between precision and recall. One can often adjust a system to retrieve more documents, thereby increasing recall, but at the expense of retrieving more irrelevant documents, and thus decreasing precision.

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.<sup>33</sup>

30 See Cyril W. Cleverdon et al., ASLIB CRANFIELD RESEARCH PROJECT: FACTORS DETERMINING THE PERFORMANCE OF INDEXING SYSTEMS (1966) (Vol. I, Design), available at [http://www-nlpir.nist.gov/projects/irlib/pubs/cranv1p1/cranv1p1\\_index/cranv1p1\\_toc.html](http://www-nlpir.nist.gov/projects/irlib/pubs/cranv1p1/cranv1p1_index/cranv1p1_toc.html); Cyril W. Cleverdon et al., ASLIB CRANFIELD RESEARCH PROJECT: REPORT OF CRANFIELD II (1966) (Vol. II, Test Results), available at [http://www-nlpir.nist.gov/projects/irlib/pubs/cranv2/cranv2\\_index/cranv2\\_toc.html](http://www-nlpir.nist.gov/projects/irlib/pubs/cranv2/cranv2_index/cranv2_toc.html); see generally, C.J. VAN RJIJSBERGEN, INFORMATION RETRIEVAL (2d ed. 1979), available at <http://www.dcs.gla.ac.uk/Keith/Preface.html>.

31 See Ricardo Baeza-Yates & Berthier Ribeiro-Nero, MODERN INFORMATION RETRIEVAL 437-455 (1999) (glossary), available at <http://www.sims.berkeley.edu/~hearts/irbook/glossary.html>.

32 *Ibid.*

33 There are many other common metrics that are considered in IR literature, including F-measure, mean average precision and average search length. F-measure is an approximation of the cross-over point between precision and recall, which allows one to see where the compromise is between the two. Mean average precision determines the existing 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-relevant items retrieved to the total number of items retrieved, and “elusion,” the proportion of responsive documents that have been missed.

### *Measuring the Efficiency of Information Retrieval Methods*

Efficiency is important to the usability of an IR system, but it does not affect the quality of the results. Efficiency is measured in two ways. The first measurement is the mean time for returning search results. This can be measured by average time to return the 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 usability of the IR system.

#### *The Blair and Maron Study*

The leading study testing recall and precision in a legal setting was conducted by David Blair and M.E. Maron in 1985.<sup>34</sup> It is a classic in showing the problem 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 informational records.

Indeed, Blair and Maron found that attorneys were only about 20% effective at thinking up all of the different ways that document authors could refer to words, ideas, or issues in their case.

The case involved a San Francisco Bay Area Rapid Transit (BART) accident 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 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, depending on their role. 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*, n.34, 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.”

#### *The Impact of Ambiguity and Variation on Precision and Recall*

Since the Blair and Maron study, some further efforts have been made to study the precision/recall issues in a legal discovery context, some of which have been performed by members of The Sedona Conference.<sup>35</sup> This field requires further study.

The limitation on search and retrieval methodology exposed in the Blair and Maron study was not the ability of the computer to find documents that met the attorneys’ search criteria, but rather the inability of the attorneys and paralegals to anticipate all of the possible ways that people could refer to the issues in the case. The richness of human language causes a severe challenge in identifying informational records.

*Ambiguity* refers to the tendency of words and expressions to have different meanings when used in different contexts. These contexts are “referential variants” or *variation*. If one and only one word or expression is found in only one and only one context, it would present no ambiguity and no

34 David C. Blair & M.E. Maron, “An evaluation of retrieval effectiveness for a full-text document-retrieval system,” *Communications of the 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.* at 225 (2007).

35 See, e.g., Anne Kershaw, “Automated Document Review Proves Its Reliability,” DIGITAL DISCOVERY & E-EVIDENCE, Nov. 2005, at 10, 10-12 (client-sponsored private study); 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 212-220 (using structured caselaw in Westlaw databases); see also Text Retrieval Conference, <http://trec.nist.gov/>, discussed *infra* Part VII.C.

variation. A search for that term would retrieve all of the documents in which the term appeared, and all of the documents would be relevant. While there may not be an exact mathematical comparison, generally speaking, the lower the variation in the contexts, the lower the likely overall recall, and the lower the ambiguity of the search term, the better the precision of the result.

But as the Blair and Maron study demonstrates, human language is highly ambiguous and full of variation. In the years since Blair and Maron, the IR community has been engaged in research and development of methods, tools, and techniques that compensate for endemic ambiguity and variation in human language, and thus maximize the recall and precision of searches.

## V. BOOLEAN AND BEYOND: A WORLD OF SEARCH METHODS, TOOLS AND TECHNIQUES

In the twenty years since the Blair and Maron study, a variety of new search tools and techniques have been introduced to help find relevant information and to help 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 function and task. It is important to know what each methodology does when it is used alone or in conjunction with other methodologies.

Clearly, different search methods have different functions and values in different circumstances. There is no one best system for all situations, a key fact for practitioners learning the technique of search and retrieval technology.

A more detailed description of search methods and techniques is set out in the Appendix. These methods can be grouped into three broad categories, but there are hybrid and cross-cutting 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”).

### *Statistical Techniques*

Second, there are a variety of *statistical techniques*, which analyze word counts (how many times the same keyword will appear in a document, or will appear near other keywords). One such approach is called “Bayesian,” derived from a famous mathematical theorem. Querying the data set using combinations of one or more of these types of Bayesian methods may well result in returning a broader slice of the data than merely using a simple keyword search, or a keyword search with Boolean operators.

### *Categorizations of Data Sets*

Third, there are other techniques depending on *categorizations of the entire data set* with various methodologies heavily reliant on setting up (*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 continually, as more data is added to the data set.

However, data sets generally need to be indexed to use any of the latter alternative 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 coded 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. Indexing a set number of structured fields (*i.e.* coded data) will be much faster because only those 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 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). As the Blair and Maron study shows, people coming in after the fact are actually very poor at guessing the right words to use in a search – words that find the documents a person is looking for without overwhelming the retrieval with irrelevant documents. In light of this fact, alternative search methods may serve to help to organize large collections of documents in ways that people have trouble doing.

Using a thesaurus, taxonomy, or ontology generally gives 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 relationships in the text that might not have been anticipated. This means that one gets unexpected results from time to time, which can be of great value, but can also be somewhat disconcerting (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 documents containing the word “Klonopin” even 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 relational contexts; the systems, however, can go help to derive them directly from the documents processed.

Finally, none of these systems is magical. Language is sometimes shared just between two people, who have invented a shorthand or code. All tools require common sense, based on a thought-out approach. Some techniques may be difficult to understand to those without technical backgrounds, but they need not be mysterious. If a vendor will not explain how a system works, it is most likely because of ignorance. Ask for someone who can provide an explanation.

There is no magic to the science of search and retrieval: only mathematics, linguistics, and hard work. If lawyers do not become conversant in this area, they risk surrendering the intellectual jurisdiction to other fields.

## VI. PRACTICAL GUIDANCE IN EVALUATING THE USE OF AUTOMATED SEARCH AND RETRIEVAL METHODS

***Practice Point 1.*** *In many settings involving electronically stored information, reliance 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.*

For the reasons articulated in prior sections, the demands placed on practitioners and parties in litigation and elsewhere increasingly dictate that serious consideration be given to 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 involved in undertaking 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 only to

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 necessary.

Even in less complex settings, sole reliance on manual review may nevertheless be an inefficient use of scarce resources. This is especially the case where automated search tools used on the front end of discovery may prove to be useful in a variety of ways, including for sampling, categorizing or grouping documents in order to facilitate later manual review.

Of course, the use of automated search methods is not intended to be mutually exclusive with manual review; indeed, in many cases, both automated and manual searches will be conducted: with initial searches by automated means to cull down a large universe of material to more manageable size, followed by a secondary manual review process. So too, while automated search methods may be used to find privileged documents out of a larger set, it remains the case that the majority of practitioners still will rely on manual review processes to identify the bases for privilege to be asserted for each document.

***Practice Point 2. Success in using 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 a discovery obligation. 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 put into action at the earliest opportunity. This knowledge can take many forms.

First, an evaluation of the legal setting a party finds itself in is of paramount importance, since the nature of the lawsuit or investigation, the field of law involved, and the specific causes of action under which a discovery obligation arises must all be taken into account. For example, keyword searches alone in highly technical patent cases may prove highly efficacious. In other types of cases, including those with broad causes of action and 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 a “relevance needs analysis,” to first 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 parties who would be the “owners” of relevant data. Time and cost considerations must also be factored in, including budgeting for human review time. These practice points apply whether your client is a defendant and holds a universe of potentially discoverable data, or your client is a plaintiff party who is expecting to receive similarly massive data in response to requests for documents.

***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.***

The choice of a search and retrieval method for a given situation depends upon a number of factors.

For example, a search method that eliminates false positive “noise” (achieving high levels of precision) may not yield the highest number of relevant documents. In other cases, such as sampling, a search method will be graded on its ability to measure statistical significance of the occurrence of a particular word or concept. 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 overall relevant universe of electronically stored information is a significant factor. Electronically stored information 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 the 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. Yet in other cases, it will be irrelevant.

Next, the volume and condition of the electronically stored information, and the extent to which electronically stored information is contained within static or dynamic electronic applications is relevant to the decisions made by the advocate or investigator.

Third, the time it will take to use a particular search and information retrieval method and its cost, as compared to other automated methods or human review, must 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 vs. finding responsive documents as efficiently as possible, *i.e.*, with the least number of nonresponsive documents). In other words, one must consider the desired trade off between recall and precision. 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 electronically stored information 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 logistical constraints of the representatives of the party making the selection (attorneys, litigation support staff, vendors).

Sixth, there is the 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 published papers supporting the reliability of the search and information retrieval method for particular types of data, or in particular settings.

***Practice Point 4. Parties should perform due diligence in choosing a particular information retrieval product or service from a vendor.***

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 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 what 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 your own 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, in terms of its flexibility. 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 de-duplication, to trace email threads, and to provide listings of related terms or synonyms, are all examples of the kind of specific functional requirements that should be inquired about.

Other types of due diligence inquiries may involve administrative matters (*e.g.*, understanding maintenance and upkeep, additional charges, system upgrades, availability of technicians, system performance), quality control issues (*e.g.*, prior testing of the method or tool in question; how databases and dictionaries supporting concept searching were populated; how strong is

the application development group of the provider), and, finally, any relevant licensing issues, involving proprietary software or escrow agreements with third parties.

**Practice Point 5.** *The use of search and information retrieval tools does not guarantee that all responsive documents will be identified in large data collections, due to characteristics of human language. Moreover, differing search methods may produce differing results, subject to a measure of statistical variation inherent in the science of information retrieval.*

Just as with past practice involving manual searches through traditional paper document collections, there is no requirement that “perfect” searches will occur – only that lawyers and parties act reasonably in the good faith performance of their discovery and legal obligations. From decades of information retrieval research, we know 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, results in all present day automated search methods falling short.

It is also important to recognize that there will 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 (such as one based on statistical properties of how words appear in the data set), 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 average measures of recall and precision), rather than being judged by whether it produces the identical document set as compared with a different technique. One possible benchmark to employ when considering use of an alternative search method is to compare the results of such a search against a similar search utilizing keywords and Boolean operators alone.

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 should make a good faith attempt to collaborate on the use of particular search and information retrieval methods, tools and protocols (including as to keywords, concepts, and other types of search parameters)*

The *Treppel* decision and other recent case law indicates that courts are becoming more comfortable with addressing search and retrieval issues, particularly in the context of blessing 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 even in advance of the 2006 amendments to the Federal Rules of Civil Procedure.<sup>36</sup>

Under newly modified Rule 26(f), the parties’ initial planning is expected to 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 easily should encompass a specific discussion on search methods and protocols to be employed by one or both parties. While disclosure of these methods and protocols is not mandated or legally required under this rule, the advantages of collaborating should strongly be considered.

<sup>36</sup> See Kenneth J. Withers, “The December 2006 Amendments to the Federal Rules of Civil Procedure,” 4 Nw. J. OF TECH. & INTELL. PROP. 171 (2006), available at <http://www.law.northwestern.edu/journals/njitip/v4/n2/3> (what “probably strikes the reader [of *Treppel*] 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]).”

Reaching an early consensus on the scope of searches has the potential to minimize the overall time, cost, and resources spent on such efforts, as well as minimizing the risk of collateral litigation challenging the reasonableness of the search method employed.<sup>37</sup>

***Practice Point 7. Parties should expect that their choice of search methodology will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and trials).***

Counsel should be prepared to explain what keywords, search protocols, and alternative search methods were used to generate a set of documents, including ones made subject to subsequent manual searches for responsiveness and privilege. This explanation may best come from a technical “IT” expert, a statistician, or an expert in search and retrieval technology. Counsel must be prepared to answer questions, and indeed, to prove the reasonableness and good faith of their methods.

***Practice Point 8. Parties and the courts should be alert to new and evolving search and information retrieval methods.***

What constitutes a reasonable search and information retrieval method is subject to change, given the rapid evolution of technology. The legal community needs to be vigilant in examining new and emerging techniques and methods which claim to 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.

## 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 information retrieval research community based on a shared interest in how to improve the accuracy and efficiency of information retrieval?

### A. Harnessing the Power of Artificial Intelligence (AI)

A statement from page 36 of The Sedona Conference<sup>®</sup>, *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 behavior all present steep challenges to the AI community in developing optimal search and retrieval techniques.

But the future continues to hold promise. Not only is there the possibility of applying sophisticated artificial intelligence means to data mining of traditional texts, but looming immediately on the horizon are new and better approaches to image and voice pattern recognition. Clearly, all forms of data stored in corporations and institutions will be fair game in terms of being within the scope of future information demands in legal settings.

Finding information on the Web sometimes is easier than finding documents on one’s own hard drive. The post-Google burst of interest in building better search engines for the Web can only

<sup>37</sup> See G. Paul and J. Baron, *Information Inflation*, *supra* n.10, at Paragraphs 50-55 (discussing an iterative collaboration process that includes adoption of multiple “meet and confers” to discuss and refine preliminary search results).

help lead to new and better search techniques applied in more well-defined contexts, such as corporate and institutional intranets and data stores.

A recent “2020 Science” report issued by Microsoft anticipates 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.”<sup>38</sup> 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. The field of future AI research in the specific domain of search and retrieval is wide open.

### **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 – a set of instructions outlining the workflow for the tool. How well these methods are applied significantly impacts the performance, and therefore, the results generated by the technology. This is where process comes in. Process functions to provide order and structure by setting guidelines and procedures designed to ensure that a technology performs as intended. Effectively applied, process will then drive 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 will then establish the technology’s credibility and value.

#### *The Important Nature of Process*

A process is a considered series of events, acts or operations leading to a result or an 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 can not 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 parties in making informed decisions.

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. All together, risk is reduced and confidence is promoted.

Search and retrieval should be visualized as a process which enables a party to distinguish potentially discoverable information from among a broader set of electronic data for purposes of production. It consists of several process 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 increases greatly. Additionally, greater volume increases the likelihood of the omission of some relevant documents. By developing and implementing process steps that consistently address these issues, their impact can be diminished and the reasonableness and good faith of the technology can be established.

---

38 See <http://research.microsoft.com/towards2020science/downloads.htm>, p 15.

### *“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 expects that every relevant, non-privileged document will be found and produced, the standard against which we measure these new technologies and processes must be based upon the same principles that have traditionally governed discovery – reasonableness and good faith. Although these terms conjure thoughts of ambiguity and uncertainty, they can actually represent a well-defined set of expectations when placed within the context of process.

A process that emphasizes reasonableness and good faith 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.<sup>39</sup> 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.<sup>40</sup> As noted in *Sedona Principles* 6 and 11, a party may choose to implement this approach in a reasonable manner, which is left to the good judgment of the party.

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 and good faith can be defined and measured by identifying performance criteria based on their attributes. Accordingly, the unreasonable and unattainable goal of “perfection” should not be allowed to be an enemy of the attainable and measurable goal of reasonableness.

As search and retrieval technologies and associated processes are developed, parties will no doubt want to use them in order to achieve defensible and credible results. If a party fails to adhere to appropriate performance guidelines it will be subject to scrutiny and criticism. Therefore, established process in conjunction with sound technology can serve as a benchmark for conducting discovery in the future. Furthermore, defensibility in court will very likely depend on the implementation of, and adherence to, processes developed for use with a search and retrieval technology.

### *Implementing Process*

Using a search and retrieval technology in conjunction with an implementing process in the complex context of electronic discovery will involve multiple phases of activity, with iterative feedback opportunities at appropriate decision points to allow integration of what a case team learns after each exercise of the process in order to calibrate and maximize the technology’s capability 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 with the intent that it serve 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 is to initiate the process by focusing on the information collected from a few of the custodians who were at the center of the facts at issue in the litigation or investigation. Focusing on information collected from the core custodians, which has a higher likelihood of being relevant, will help the team efficiently develop its understanding of the issues and language used by the custodians, thus allowing them to more efficiently develop and implement an appropriate search and retrieval process.

---

<sup>39</sup> *Hickman v. Taylor*, 329 U.S. at 507; see *supra*, n.1.

<sup>40</sup> Under 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.”

The initial selection and refinement of search terms can also benefit from the application of sampling techniques that can help the review team to rank the precision and recall 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 the review team to focus the selection of terms.<sup>41</sup>

The development of process control logs and second-level review techniques can also help the review team to ensure that the designed process is consistently applied to all of the information to be reviewed. Additionally, a second-level review process based on statistical sampling techniques can ensure the achievement of 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.

### C. 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 growing consensus that the application of linguistic and mathematic-based content analysis, search and retrieval technologies, and tools, techniques and process in support of the review function can effectively reduce the cost, time, and error rates.

#### Recommendations

1. *The legal community should 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, for use in assisting lawyers in evaluating the competitive legal and regulatory search and retrieval services market.*

As stated, in the 20 years since the Blair and Maron study, there has been little in the way of peer-reviewable research establishing the efficacy of various methods of automated content analysis, search, and retrieval as applied to a legal discovery context. A program of research into the relative efficacy of search and retrieval methods would 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. Moreover, the outcome of the application of advanced content analysis, search and retrieval methods can have significant differences based on expertise of the operator. Ideally, a research program would advance the goals of setting minimum or baseline standards for what constitutes adequate information retrieval, as well as reaching agreement on how to benchmark competing methods against agreed-upon objective evaluation measures.

In this regard, The Sedona Conference<sup>®</sup> supported the introduction of a new “Legal Track” in 2006 for the TREC research program run by the National Institute of Standards and Technology. NIST is a federal technology agency that works with industry to develop and apply technology, measurements and standards. TREC is designed “to encourage research in information retrieval from large text collections.”<sup>42</sup> The TREC legal track involves an evaluation of a set of search methodologies

<sup>41</sup> See text at Part IV, *supra*.

<sup>42</sup> The Text Retrieval Conference (TREC) was started in 1992. See <http://trec.nist.gov>. Its purpose is to support research within the information retrieval 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 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 that is a forum for participants to share their experiences. The TREC test collections and evaluation software are available to the retrieval 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 information retrieval and of facilitating technology transfer, and many of today's commercial search engines include technology first developed in TREC.



based on lawyer relevancy assessments on topics drawn from a large public database of OCR-ed documents. The results coming out of the 2006 legal track represent the type of objective research study into the relative efficacy of Boolean and other search methods that the legal community should further encourage.<sup>43</sup>

However, a need exists to scale up the TREC research to accommodate the potential retrieval of millions or tens or hundreds of millions of arguably relevant documents among a greater universe of terabytes, petabytes, exabytes, and beyond.

Members of The Sedona Conference<sup>®</sup> community have and will continue to participate in collaborative workshops and other fora focused on issues involving information retrieval.<sup>44</sup> How best to leverage the work of the IR community to date is an enterprise beyond the scope of this paper. The Sedona Conference<sup>®</sup> 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.

---

<sup>43</sup> See Jason R. Baron, David D. Lewis & Douglas W. Oard, "TREC 2006 Legal Track Overview," 2006 FIFTEENTH TEXT RETRIEVAL CONFERENCE (TREC 2006) PROCEEDINGS, available at <http://trec.nist.gov/pubs/trec15/papers/LEGAL06.OVERVIEW.pdf>; see also TREC 2007 Legal Track, <http://trec-legal.umiacs.umd.edu/> (additional documentation relating to TREC 2006 Legal Track).

<sup>44</sup> See, e.g., *Supporting Search and Sensemaking for Electronically Stored Information in Discovery Proceedings* ("DESI Workshop"), held at the Eleventh International Conference on Artificial Intelligence and Law (ICAIL 2007), June 4, 2007, Palo Alto, CA, papers available at <http://www.umiacs.umd.edu/~oard/desi-ws/>.

## APPENDIX: Types of Search Methods

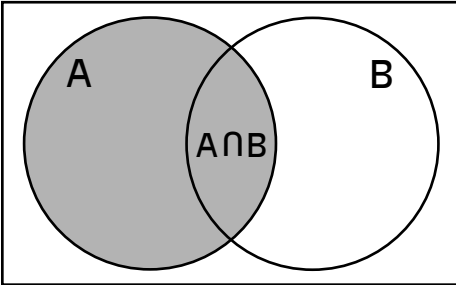
*This appendix is a “survey” of different forms of search methods found in the information science literature, and which form the basis of offerings by vendors in the legal marketplace. The list is not definitive. Indeed, as the main body of the Commentary makes clear, rapid technological progress will inevitably affect how methods are described, perfected, and then replaced with new ways of performing search and retrieval.*

*A second caveat: the following search methods are not intended to be mutually exclusive. Indeed, many products tout the benefits of hybrid, combined, or cumulative approaches to performing searches.*

### A. Boolean Search Models

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 information retrieval as computers became more widely accepted. Boolean search statements can easily be applied to large sets of unstructured data and the results exactly match the search terms and logical constraints applied by the operators.

As used in set theory, a Boolean notation demonstrates the relationship between the sets or groups, indicating what is in each set alone (set union), what is jointly contained in both (set intersection), and what is contained in neither (set differences). The operators of AND (intersection or  $\cap$ ), OR (union or  $\cup$ ) and AND NOT or BUT NOT (difference) are the primary operations of Boolean logic. These relationships can easily be seen within a Venn diagram (see below).



**OR** is a Boolean operator that states 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  $A \cup B$  (the entire shaded areas above). The use of OR expands the resulting Boolean set.

**AND** is a Boolean operator used to identify the intersection of two sets or two keywords. The purpose of this command is to help construct more complex concepts from more simple vocabulary words. AND is represented by the middle intersecting area above ( $A \cap B$ ). The use of AND restricts the resulting Boolean set.

**NOT** is a Boolean operator used to eliminate unwanted terms. The purpose of this command (preceded by either AND or BUT) is to help suppress multiple meanings of the same term; in other words, eliminating ambiguity.

Different search engines or search tools may provide additional Boolean-type 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 logical 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 sometimes can lead to ambiguous cases. 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:** Another extension to Boolean searching, this technique checks the position of terms and only matches those within the specified distance. This is a useful method for establishing relevancy between search criteria, as well as for paring down irrelevant matches and getting better results (improving precision). Some search engines let you define the order, in addition to the distance. For example: *budget w/10 deficit* might mean “deficit within the 10 words following the word budget”.
- **Phrase Searching:** Some search engines provide an option to search a set of words as a phrase, either by typing in quote marks (“ ”) or by using a command. When they receive this kind of search, the engines will generally locate all words that 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 and stemming). This search capability allows the user to widen the search by searching a word stem or incomplete term. It is typically a symbol such as a question mark (?), asterisk (\*), or exclamation point (!). The search system 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 systems will allow for internal truncation such as wom?n would find women or woman. The \* and ! terms have broader application: for example, hous\* would find house, housemate, Houston, household or other similar words with the stem “hous.”

## B. Probabilistic Search Models: Bayesian Classifiers

Probability theories are used in information retrieval to make decisions regarding relevant documents. The most prominent of these are so-called “Bayesian” systems or methods, based on Bayes’ Theorem. The theorem was developed by Thomas Bayes, an eighteenth century British mathematician. A Bayesian system sets up 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 it appears, the more weight it carries.
- Closer to the top – documents with the term in the title are more heavily weighted
- Adjacency or proximity – the closer the terms are to each other, the higher the weighting
- Explicit or implicit feedback on relevance

(Note: other types of search models apply these types of concepts or ideas as well.)

Bayesian systems frequently utilize a “training set” of highly relevant documents to increase understanding, and therefore the probability measures of the system. During training, the system examines each word in the training documents and computes the probability with which that word occurs in each category.

For example, the word “potato” may occur in 5 documents in the category “kitchen tools” (*e.g.*, “potato peeler”), in 7 documents in the category “farm products,” and in one document in the category “garden tools.” When a new document is then found to contain the word “potato,” the Bayesian classifier will interpret this new document as most likely to be a member of the category “farm products” than either of the other two. The same process is repeated for all of the words in the document. Each word in the document provides evidence for which of the categories the document belongs to. The Bayesian classifier combines all of this evidence, using Bayes’ rule, and determines the most likely category.

Bayesian classifiers provide powerful tools for comparing documents and organizing documents into useful categories with a moderate amount of effort.

### C. Fuzzy Search Models

Boolean and probabilistic search models rely on exact word matches to form the results to a query. Exact matching is very strict: either a word matches or it doesn’t. Fuzzy search is an attempt to improve search 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 related to stemming in Boolean classifiers, discussed above.

Some algorithms for fuzzy matching use 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 the matching letters in the center than at the edges. Unfortunately, this can sometimes bring up results that make little sense (a search for *tivoli* might bring up *ravioli*).

Many systems allow one to assign a degree of “fuzziness” based on the percentage of characters that are different. Fuzzy searching, or matching, has at least two different variations: finding one or more matching strings of a text, and finding similar strings within a fixed string set often referred to as a dictionary. Fuzzy searching has many applications in legal information retrieval including: spellchecking, email addresses and OCR clean-up.

### D. Statistical Methods: Clustering

Systems may use statistics or other machine-learning tools to recognize what category certain information belongs to. 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) 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.

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.

### E. Machine Learning Approaches to Semantic Representation

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 embrace 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 the 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 their 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 almost as satisfied to find a document that did not contain that exact word, but did contain one of these related words.

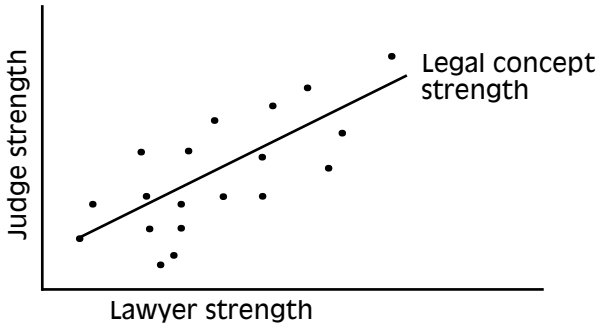


Figure 1. Dimension reduction – the original dimensions of “lawyer” and “judge” are combined into a single dimension. Each point in the graph represents a document. Its location in the graph shows how much the document is related to each dimension.

The figure above illustrates the kind of relationships such systems find. 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 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 the 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. And, 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,<sup>45</sup> we can describe the document by how it relates to each of these reduced dimensions. Latent Semantic Indexing (LSI, also called Latent Semantic Analysis) is probably the best known of these dimension-reducing techniques, but there are others, including neural networks and other kinds of statistical language modeling.

Such 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 (say 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.

45 See H. Roitblat, *supra*, n.34.

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.

### 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 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 kinds of law personnel. The only relations typically included in a taxonomy are 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” and “Motion Picture and Sound Recording Industries” and “Broadcasting.”

One can use this kind of taxonomy to recognize the conceptual relationship among these different types of personnel. If your 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.

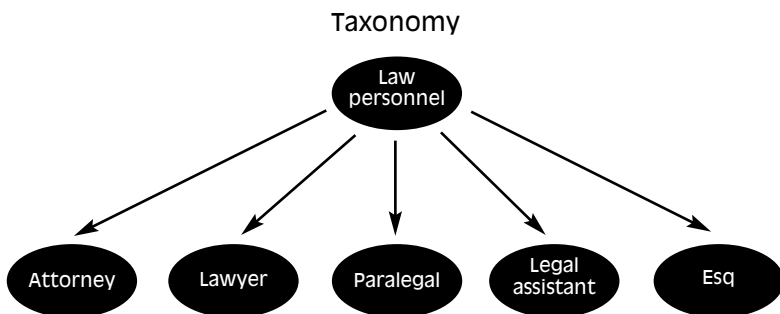


Figure 2. A simple taxonomy for law personnel.

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 your category includes attorneys, you 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.

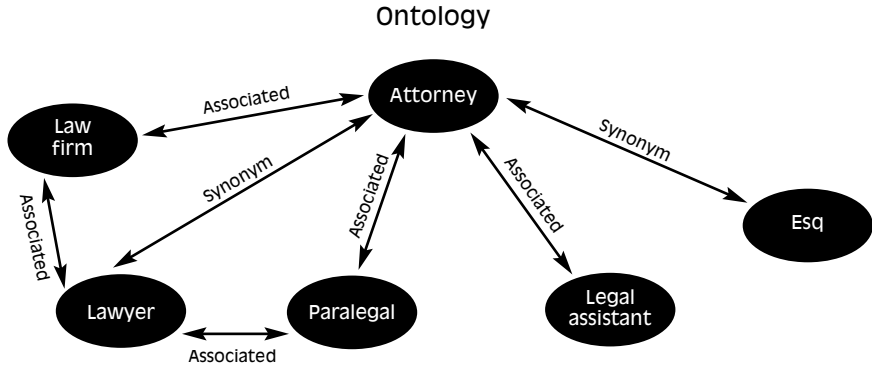


Figure 3. A section of an ontology of legal personnel.

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 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 researcher can more efficiently study the research 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. Table – 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 multi-dimensional pivot table; includes, table, group, timeline and tree functionality

In practice, a researcher can load search results into a presentation technology for an organized view, and then drill-down to access discrete items of significant interest or concern. This often iterative process may help a researcher to learn more about, act on, and manage search results.



# SEARCH AND INFORMATION RETRIEVAL SCIENCE

---

*Herbert L. Roitblat, Ph.D.*  
*Orcatec LLC*  
*Ojai, CA*

## I. A BRIEF HISTORY OF INFORMATION RETRIEVAL

The problem of how to use automated systems to find stored information is at least as old as computers themselves. In 1945, Vannevar Bush, Director of the Office of Scientific Research and Development, noted:

The difficulty seems to be, not so much that we publish unduly in view of the extent and variety of present day interests, but rather that publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships (The Atlantic, June 1945<sup>1</sup>).

Bush pointed to electronic, photographic and other emerging “mechanical aids with which to effect a transformation in scientific records.” Even Bush’s prophetic notions of how machines would make information storage and retrieval more effective grossly underestimated the difficulty of indexing and retrieving that information. The development of general purpose computers whose capabilities far surpassed those of the special purpose machines that Bush had in mind and the proliferation of machine readable information caused the information retrieval industry to explode.

Several mechanical systems, like those that Bush described, were in use in the 1940s and 1950s, some even later. Some systems used punch cards and rods. Each of the collection’s topics was represented by a hole around the edge of a card. If a document was relevant to that topic, then the edge of its hole would be torn out. A knitting needle or other rod would be inserted into the deck and the whole collection would then be shaken. The relevant cards would fall out of the deck. Obviously, systems like these were very limited in the number of documents and number of topics they could handle.

The “Rapid Selector”<sup>2</sup> was another device that stored information on reels of microfilm along with machine-readable index marks. It was capable of storing 72,000 frames of information on 2000-foot film reels. Each frame consisted of a picture of a document page and an array of dots indicating the topics for which the page was relevant. At least this is the way it was supposed to work. According to some reports, the mechanics of the system were never quite up to the promise. Later information retrieval systems continued to use similar index systems, in which each document is indexed by one or a few specific terms derived from preconstructed classification schemes and subject heading catalogs (The Library of Congress cataloging scheme is an example of such a catalog system for books). Thesauruses and other hierarchical vocabulary schemes made it possible to expand the

---

<sup>1</sup> <http://www.theatlantic.com/doc/194507/bush>

<sup>2</sup> <http://www.ischool.berkeley.edu/~buckland/goldbush.html>

vocabulary by adding more specific or more generic concepts than those originally specified. These systems are called controlled vocabulary systems because documents were only indexed by a few terms drawn from the hierarchy.

By the early 1950s information retrieval systems were moving away from the hierarchical catalogs to schemes based on single term descriptors or key words, coded without context. This innovation freed the searcher somewhat from the tyranny of the cataloger and allowed one to search for combinations of these single terms to produce complex queries. Searchers could use phrases and chains of synonyms. Researchers found that these coordinate keyword indexing methods were much easier to use than the hierarchical catalogs, and were no less accurate. They still relied on a controlled vocabulary of key terms, assigned by trained experts, but they could be combined in novel, unanticipated ways when searching the catalog.

Even after the introduction of commercial computers for information retrieval in the mid 1950s, indexers continued to rely on document descriptors that were drawn from a controlled vocabulary and were assigned by domain-experts. Computer memory (both internal core memory and external memory devices) was expensive, and indexing every word in the document would take up more storage than the documents themselves. Further, the documents to be retrieved were generally not machine readable, and, before OCR (optical character recognition), would have to be typed in by hand. These factors conspired to limit the range of terms that could be included in the search vocabulary. As a consequence, expertise was needed to catalog documents and further expertise was needed to find them again. Document ranking was not required to be too precise because the available computer resources at the time limited most searches to batch processing. A query would be submitted and some time later (from minutes to days) a printout of all the search results would be returned to the user.

### Text analysis

It was generally believed in the 1950s that indexing documents, that is, assigning keywords to them, was a job that could only be done well by professionals. In this context, H. P. Luhn<sup>3</sup> introduced the idea that computers could not only handle the keyword matching and sorting task, but they could actually be used to analyze the content of written texts. He proposed automatic indexing and term weighting techniques based on the frequency and location of the words in the text. For example, a term that appeared twice in a paragraph or in two succeeding paragraphs could be considered a major concept in the document. His system would then use these terms to index the document. This idea was a major milestone in computerized information retrieval. Luhn is also credited with the idea that documents could be represented mathematically by term vectors.

Typically, each position on the list represents one word. If the word is present in a text, the corresponding element in the vector representing that text is set to be nonzero. If the word is absent, the corresponding element is set to 0. So, for example, if the 518th element of a vector represents the word "collusion" it is set to be 0 unless the word "collusion" actually happens to be in the text.

In a 1960 paper, Calvin Mooers<sup>4</sup> recognized that manually assigning descriptors to large volumes of documents was impractical. The job would have to be done by machine. He proposed an "inductive inference" machine, which takes as input a set of solved examples, such as a set of documents and the corresponding set of index terms and then uses an inference mechanism to derive the rules by which those index terms were assigned.

An inductive inference machine is one that can be taught a series of correctly solved examples of problems so that it can proceed on its own (with some supervision and corrective intervention, probably) to the solution of other problems in the same class... It is capable of learning a great variety of tasks. (Mooers, 1960, p. 232)

<sup>3</sup> *E.g.*, Luhn, H.P. (1957). "A statistical approach to mechanized encoding and searching of library information", IBM J. Res. Dev., 1, no.4, 309--17.

<sup>4</sup> C.N. Mooers, "The Next Twenty Years in Information Retrieval", American Documentation, 11:3, July 1960, 229-236.





Several techniques have evolved to deal with these problems that derive from people's flexible language use. One of these, restricting the search vocabulary, has already been discussed. Whatever specific words were used in the document, the cataloger coerces the document into one or a few categories from the controlled vocabulary. Similarly, the terms a searcher has in mind have to be coerced, perhaps with the use of a thesaurus, into the same controlled vocabulary terms used by the cataloger.

Controlled vocabulary systems may work well with professionally trained search intermediaries, but they are unacceptable to a the broader search audience. Generally, free text indexing has been found to be no worse and often better than controlled vocabulary approaches. In any case, the process of classifying documents into a small number of categories is very time consuming and highly dependent on the skill of the cataloger.

Other approaches have involved augmenting the original search with additional search terms (called query expansion) or translation schemes that convert natural language queries into formal search queries or into preconstructed and edited search queries (e.g., Ask Jeeves, now Ask.com).

Query expansion converts the term or terms that the searcher enters into a more elaborate query that usually involves the original terms and some additional terms. Query expansion can be based on morphological characteristics of the words, word association patterns, thesauri, taxonomies or ontologies.

The morphology of a word is its semantic structure. Words consist of one or more morphemes. A morpheme is the smallest unit of meaning in the language. The word "unfriendly," for example, consists of three morphemes, "un" meaning "not," "fiend," meaning "companion," and "ly," which indicates that it is an adverb. Morphological expansion is often useful, in that a system that "knows about" government should also know about "govern" and "governing." In English, stemming is a simple form of morphological analysis. In stemming, the main stem of a word is separated from its inflections. For example, "swimming" consists of a stem, "swim" and an inflection or suffix, "ing." When stemming is used, a search for "swim" or "swimming" will yield the same results, all documents containing either of these word forms. Stemming does sometimes lead to inappropriate expansions (e.g., expanding a query for "Miller" to include "Mills," "Mill," "Milled" and "Milling"), but there are techniques for mitigating such difficulties.

Thesaurus-based expansion depends on the existence of an appropriate and pertinent thesaurus. Some systems employ, for example, semantically encoded dictionaries such as the Longman Dictionary of Contemporary English or George Miller's WordNet<sup>6</sup>, a manually constructed thesaurus based on psycholinguistic assessments of lexical relations. Thesauri such as these are difficult to use, because they include multiple senses and multiple meanings for each word. It can be difficult to select the intended meaning from a brief query and it may actually decrease the accuracy of some queries. In some specialized areas with their own distinct dialects, a general thesaurus can greatly interfere with correct selection. For example, in some collections a search for "cleric" might be expanded via the thesaurus to "minister." This expansion would probably be useful in the US, but would give erroneous results in documents about the Middle East, where a minister is a member of the government, and cleric is more tightly related to "Ayatollah." Solving these problems with thesauri can require a great deal of human editing, and they are very brittle—they cannot adapt easily to changes in the language community. For example, it would take considerable expert human intervention to be able to respond to new terms such the emergence of the MP3 file standard, or podcasting.

Similarly, the use of simple co-occurrence patterns, even in text of known relevance, is less than completely satisfactory because these patterns do not capture the overall regularities in the language usage. Equally associated terms are not always equally pertinent. For example, two terms that co-occur frequently in a document may have a different value than two terms that occur only rarely in the document, but always together. These simple mechanisms fail to take redundancy into account.

---

6 Fellbaum, C. (1998). *Wordnet: An electronic lexical database*. Cambridge, MA: MIT Press.

Some mechanism (either human intervention or some learning/inference system, such as a neural network) is necessary to properly weight the co-occurring terms.

Weighting terms or even pairs of terms indiscriminately, has been found to be inadequate for useful searches. Some form of differential weighting is used by practically every modern search system. For example, most information retrieval systems employ some form of term-frequency weighting—terms that occur more often in a document are considered to be more important (all other things equal) to describing the document's content than terms that occur fewer times. Similarly, terms that occur in fewer documents are considered to be more important (again, all other things equal) than words that occur in many different documents. Although term weighting usually helps to improve performance, no system is yet perfect in retrieving only and all relevant documents. Neural networks can introduce more flexible weighting schemes employing nonlinearities in the weights that vastly improve retrieval precision.

## V. SYNTACTIC TECHNIQUES

Another approach to identifying what is important in a document is to do a computerized analysis of its syntactic structure. The vector space model discards word order. Syntactic approaches, in contrast depend strongly on word order and other cues to parse each sentence into a tree structure. The topics of a text are usually contained in its nouns, so identifying the nouns in a text will go a long way toward identifying its most meaningful and characteristic components.

Names, of people, places, organizations, and so on, can also be important carriers of information. Although less than perfect, systems exist that attempt to identify that the word "Bush" is a name in the sentence "President Bush declared Louisiana a disaster area," but not a name in the sentence "He sat in the place that, before the hurricane, was occupied by a blueberry bush."

Some systems are capable of recognizing that "Bush" and "President" are not just names, they refer to the same person or entity. These systems can be very useful for retrieving information where a person could be referred to by his name, his title, the name of his house, and so on. They create a kind of thesaurus that describes the named people, places, and organizations they have detected in the different ways found in the document collection. Like a conventional thesaurus, this entity thesaurus can also be used to expand queries.

Another related source of document information exploits its overall structure. For example, the title of a document is often more valuable in identifying what it is about than any other paragraph in it. In emails it is usually easy to identify who the sender and who the recipient was, but in memos, it may not be so easy for a computer to identify this information.

## VI. USER INTERFACE

Systems differ substantially in how they interact with the user. Traditional systems tended to use simple queries that consist of one or a few words. Later systems allowed users to enter complex Boolean expressions as queries (e.g., "(air or water) w/5 (pollution and controls)"). Boolean systems allow users to specify the logical relations between query terms (such as OR, AND, and NOT). Many allow nested input.

These Boolean expressions can become quite complex and creating them is, for many, a formidable task. Many modern systems allow users to enter natural language queries such as:

What factors are important in determining what constitutes a vessel for purposes of determining liability of a vessel owner for injuries to a seaman under the Jones Act (46 USC 688)? (Turtle, 1994, p. 213)<sup>7</sup>

<sup>7</sup> Turtle, H. (1994). "Natural language vs. Boolean query evaluation: a comparison of retrieval performance." Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval. Pp. 212-220. New York: Springer Verlag.

These queries are not constrained to any particular format. Turtle (1994) found that using natural language queries like this were more effective at retrieving case law documents from Westlaw than were more traditional Boolean expressions, like

(741 +3 824) FACTOR ELEMENT STATUS FACT /P VESSEL SHIP BOAT  
/P (46 +3 688) "Jones ACT" /P INJUR! /S SEAMAN CREWMAN WORKER  
(Turtle, 1994, p. 213)

Other systems do not rely on queries from the user to organize the documents in their collection. They may use various kinds of diagrams or maps to display groups of documents and allow the user to select groups of documents for further investigation.

## VII. INFORMATION RETRIEVAL IN E-DISCOVERY

Tools like those described above have been widely applied in electronic discovery, but there are very few published studies that discuss their effectiveness.

Blair and Maron (1985) found that attorneys were only about 20% effective at thinking up all of the different ways that the document authors could refer to issues in their case. The case involved a San Francisco Bay Area Rapid Transit accident in which a computerized BART train failed to stop at the end of the line. There were about 350,000 pages in about 40,000 documents for the case.<sup>8</sup> The attorneys worked with experienced paralegal search specialists 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 of this study found that the different parties in the case used different words, depending on their role. The parties on the BART side of the case referred to "the unfortunate incident," but parties on the victim's side called it an "accident" or a "disaster." Other documents referred to the "event," "incident," "situation," "problem," or "difficulty." Proper names were often not mentioned. The limitation in this study was not the ability of the computer to find documents that met the attorneys' search criteria, but the inability of the attorneys and paralegals to anticipate all of the possible ways that people could refer to the issues in the case.

Concerning one issue, the attorneys in the case identified three terms that they thought would be adequate to retrieve relevant documents, Blair and Maron found 26 more. The original three words could not by themselves be used effectively to find relevant documents, because they retrieved too many irrelevant documents. Other search terms were needed to limit the range of documents that were returned, but this limitation came at the cost of missing documents that did not happen to have these additional terms. Coming up with the right combination of terms to yield relevant results and no irrelevant results is nearly impossible.

They 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.

## VIII. INFORMATION RETRIEVAL MEASURES

Standard information retrieval measures are precision and recall. Precision is the proportion of retrieved documents that are relevant to the query or topic. Recall is the proportion of responsive documents that have been retrieved.

$$Precision = \frac{n_{Responsive\_Retrieved}}{n_{Retrieved}}$$

$$Recall = \frac{n_{Responsive\_Retrieved}}{n_{Retrieved}}$$

<sup>8</sup> Blair D. C. & Maron, M. E. (1985). "An evaluation of retrieval effectiveness for a full-text document-retrieval system," Communications of the ACM, 28, 289-299.

If a collection of documents contains, for example, 1000 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 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  $90/100 = 0.9$ . In practice there is usually a trade off between precision and recall. One can often adjust a system to retrieve more documents, thereby increasing recall, but at the expense of retrieving more irrelevant documents, and thus decreasing precision. A query for “gold or silver,” for example, will usually return more documents about metals than a query just for “gold,” but may also retrieve documents about “gold medals” and “gold standards” as well. Metaphorically, 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.

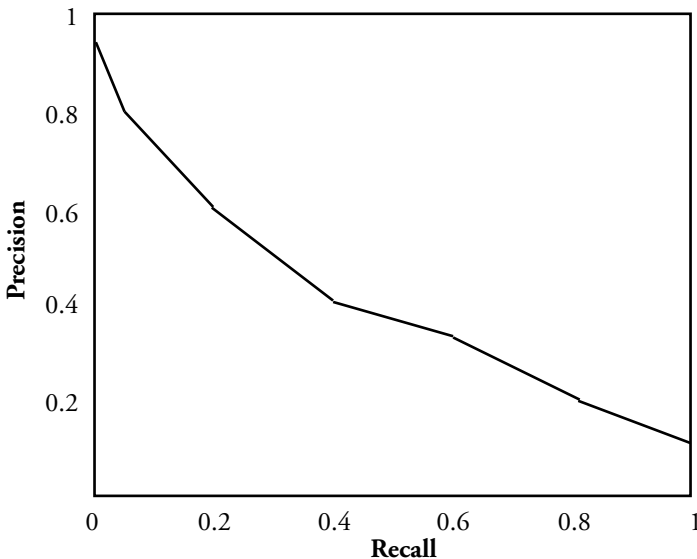


Figure 3. An example of the trade off between precision and recall.

An example of this trade off is shown in Figure 3. As more documents are retrieved, recall increases, but precision decreases. At the point at which 40% of the relevant documents had been examined, only about 43% of the examined documents would have been found to be relevant. This kind of tradeoff can be observed when you adjust the system to be more or less choosy. It is also the kind of pattern you would expect if the system returned a ranked list of documents, ranked by the probability, for example, that the document would be considered relevant, and reviewers examined each document in order. Generally, the more documents you retrieve, the higher recall will be and the lower precision will be.

Systems with the most power are those that yield higher levels of recall for a given level of precision or higher levels of precision for a given level of recall. Because the accuracy of the system is expressed as two numbers, it is often difficult to tell whether one system is better than another if, for example, it yields slightly higher recall and slightly lower precision than another system. Information retrieval investigators have developed a number of measures that combine these two into a single number. One of these is called F1 or van Rijsbergen's F. It is the harmonic mean of precision and recall. Other measures average precision obtained at different levels of recall.



In addition to the system's accuracy, there are other factors to consider when choosing among information retrieval systems. van Rijsbergen<sup>9</sup> suggests the following factors be considered when comparing systems:

- Coverage-how much relevant material can the system access?
- Speed-how quickly does the system return results?
- Effectiveness of the output-does the system present the results in a usable manner?
- Effort-how much trouble is it to set up and use the system?

In addition, evaluations might want to consider these factors:

- Usefulness-does the document provide information that is of value to the user?
- Cost-how much does it cost to get the information?
- System boundaries-to what extent does the system rely on the skill of the user to achieve high levels of performance?

In practical terms, the performance of a system is not just a matter of how well it responds to specific queries. Systems can differ in their coverage. Some systems are better able to extract usable information from a wider variety of document types than others are. Especially in large collections, the speed with which the system operates can be an important factor. The time it takes to respond to a particular query is only one of the speed measures that might be considered. This time might affect the productivity of the system.

Even if a system properly distinguishes between responsive and nonresponsive documents, how it presents this information to the user is also important. Systems differ in the effort it takes to act on the information retrieved. Some systems are just easier to use than others. Some fit with your workflow patterns better than others. Some system have more panache than others. They may look great, but actually be difficult to use.

Documents with the same relevance may differ substantially in their usefulness. Finding the 23rd copy of a significant email, for example, is much less useful than finding the first copy. Cost is almost always a consideration in electronic discovery. Can the cost of finding a few more documents be justified by the value of the documents that are retrieved? Two systems may differ only marginally in their effectiveness. Can you justify the extra expense of using the better system by the quality of the results you will receive?

System performance does not depend only on the mechanical or computational characteristics of the system, but also on the skill and knowledge of the user. The quality of the retrieved document set may depend heavily on the skill of the person formulating the queries. The role of these skills in determining the results should be evaluated along with other measures of system performance.

## IX. ALTERNATIVES TO PRECISION AND RECALL

Precision and recall are set measures. They assume that the system either returns a document (in the set) or it does not (out of the set). They also assume that a document is either relevant or it is not relevant. Many modern information retrieval systems, in contrast, rank documents by degree of relevance. Some documents are determined by the rules of the system to be more relevant to the query than others, and these are typically returned with higher ranks than the less relevant documents.<sup>10</sup>

---

<sup>9</sup> Rijsbergen, C. van (1979). *Information Retrieval*, London: Butterworths.

Figure 4 shows a useful way to think about the effectiveness of information retrieval tools that rank documents by degree of relevance or by probability of relevance. It measures both the ability to find responsive documents and to rank them in a useful way. Precision and recall only measure the ability to find responsive documents.<sup>10</sup>

The curve, called the retrieval operating characteristic (ROC) curve, represents the cumulative proportion of relevant vs. irrelevant documents retrieved at each rank in the list. An ideal system would rank all of the responsive documents before any of the nonresponsive documents. This would appear in the figure as a vertical line that goes straight up the left axis and then bends to go across the top of the graph, passing through the upper left-hand corner of the graph.. Except on trivial problems, few real systems can achieve this level of performance. Real systems tend to mix responsive and nonresponsive documents, but the better system is the one that is more likely to present the relevant ones before the irrelevant ones.

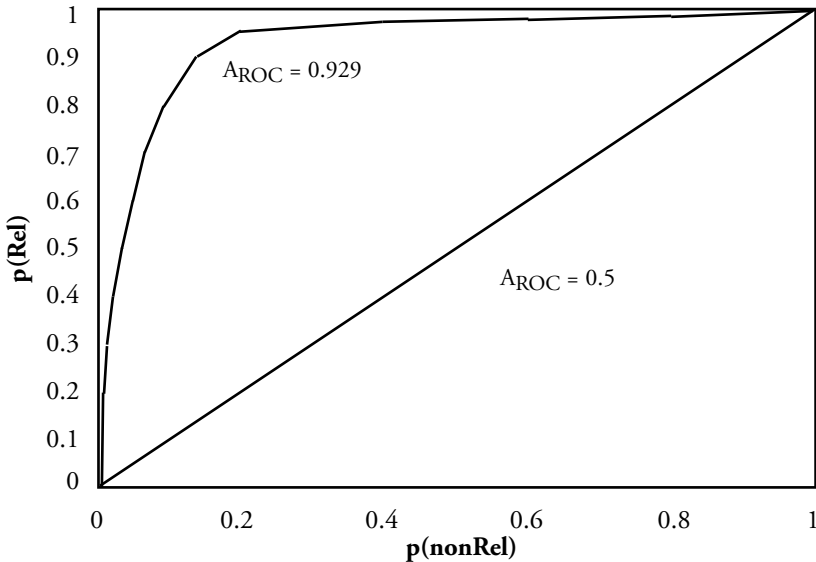


Figure 4. The Retrieval Operating Characteristic (ROC) curve also shows the trade off between precision and recall, but in an easy to summarize way. The accuracy of the system is characterized by the area under the curve ( $\text{AROC}$ )."

Precision and recall represent the accuracy of the system by two numbers. It is difficult to tell, therefore, whether a system with slightly lower recall but higher precision is more accurate than one with slightly higher recall and lower precision. There are a number of conventional ways to combine these two measures, but none of these is entirely satisfactory. All of them seem to be derived more from convention than from a deep theoretical perspective. To be sure, they are important in comparing one system to another, but they are less useful in assessing the absolute effectiveness of a system. They also do not typically take into account the goals of the person using the system. For example, the harmonic mean of precision and recall is one way of summarizing the performance of a system with fixed parameters, but it is difficult to interpret. Similarly, calculating precision at a fixed percentage of recall may be useful if your goal is to examine the top documents until you can answer a specific question. Average precision at 10%-increments in recall (*i.e.*, 0%, 10%, 20% ...90%, 100% recall) may be useful if you are interested in the overall accuracy of the system, but it is difficult to interpret. Is 40% average precision good or poor performance? These measures also depend to varying degrees on the percentage of documents in a collection that are relevant and on the stringency of the person or system making the decision.

<sup>10</sup> Swets, J.A., (1963) "Information retrieval systems", Science, 141, 245-250. See discussion in Rijsbergen, C. van (1979). *Information Retrieval*, London: Butterworths.

Similar questions have arisen in medical diagnosis. There, the task is to distinguish between those individuals who are ill with a particular disease and those who are not. Medical tests have been measured in terms of sensitivity (the ability to find ill people) and specificity (the ability to distinguish well from ill). Medical sensitivity corresponds to recall in information retrieval and medical specificity corresponds to precision. One could make the test lax so that more people are recognized as disease holders, or more stringent, so that fewer people are designated disease holders. For example, is there an epidemic of obesity (depends on how fat a person has to be to be called obese)? Is there an epidemic of autism (depends on the tests used to diagnose autism).

Although sensitivity and specificity measures continue to receive widespread use, some medical investigators use measures based on signal detection theory.<sup>11</sup> Similar measures have also received attention in information retrieval studies.

As an alternative to using two numbers—precision and recall—to summarize retrieval performance, you can summarize the accuracy of a system by calculating the area under its retrieval operating curve (ROC). The ROC incorporates the idea that items vary in the degree to which they provide evidence for a document being relevant. Some documents are either more relevant than others, or are simply more likely to be relevant than others.

ROC analysis also recognizes that the performance of a decision system, such as whether or to what degree a document is relevant, depends on two questions. To what degree are the two classes of documents separable from one another and how selective do I want to be in accepting documents as being relevant? If the two classes were identical, then, of course, no system would be able to tell them apart. The better system is the one that is better at distinguishing relevant from irrelevant documents. On the other hand, you could get higher or lower levels of precision or recall, by changing the stringency of your decision. You could require more evidence for a document to be called relevant or you could be more accepting and call a document relevant based on very little evidence. How strict or lax your criterion is depends on the task you are trying to perform and the costs of various kinds of errors, but it does not change the overall power of the system to distinguish between relevant and irrelevant documents.

In an ROC analysis, a random system that cannot distinguish between relevant and irrelevant documents will produce an ROC that is a straight diagonal line. The probability of a document being relevant or irrelevant is the same. A system that distinguishes perfectly between relevant and irrelevant documents will have an ROC that ranks all of the relevant documents before any of the irrelevant ones. Its ROC would travel up the vertical axis to the top and then travel across the top of the graph.

The area under this curve ( $A_{ROC}$ ) is a good way to summarize the performance of this system independently of your criterion. An ideal system will have an area of 1.0. A totally ineffective system, one whose ranking is unrelated to the relevance of the documents, will have an area of 0.5. This AROC measure has the advantage of characterizing a system's accuracy by a single number. It recognizes explicitly that there is a tradeoff between recall and precision that is not related to the power of the system to discriminate between relevant and irrelevant documents. One can place the cutoff between retrieved and nonretrieved documents anywhere along a system's retrieval operating curve, thereby changing its precision and recall without changing its retrieval effectiveness at all. Finally, the  $A_{ROC}$  has a straightforward absolute interpretation. Areas near 0.5 reflect poor performance whereas areas near 1.0 reflect excellent performance.

## X. ESTIMATING RECALL

Measuring system performance by either of these approaches typically requires a rather substantial effort. Precision and recall measures are designed to assess the degree to which the retrieval of responsive documents is accurate and complete. Of the two, precision is relatively easy to measure

---

<sup>11</sup> See John A. Swets and Ronald M. Pickett, *Evaluation of diagnostic systems: methods from signal detection theory*, Academic Press, New York, 1982.

because one has only to assess the responsiveness of those documents that were retrieved. Recall, on the other hand, is much more difficult to measure because one has to know how many documents were actually responsive in the whole collection in order to know the proportion of those actually responsive documents that were retrieved. This is practical only in relatively small data sets because it requires that every document be assessed for responsiveness. Sampling could be used to get an estimate of the total number of responsive documents, but it presents some challenges. Further, it is not immediately obvious what you can do with this information once it has been obtained.

One approach to estimating recall is to take a random sample of documents (without regard to whether they have been retrieved or not) and evaluate this random sample for responsiveness. Once a reasonably sized sample of responsive documents has been obtained by this method, it is a simple matter to count the proportion of those documents that have been retrieved. The number of responsive documents that must be found using this random search procedure is

$$n = \frac{Z^2 p(1-p)}{C^2}$$

Where  $Z$  is the confidence level in units of the normal distribution and  $C$  is the confidence interval. The confidence level is the overall confidence you want to have in the quality of the results. Confidence levels of 95% to 98% are typical for most situations. We will use 98%, meaning that you are 98% confident in the outcome of the measure. Higher levels of confidence can only be achieved with much greater effort. By convention, statisticians often talk about the complement of confidence or  $\alpha$  ( $\alpha = 1.0 - \text{confidence}$ , 0.02 to 0.05 in our examples). The confidence interval is how precise you want your estimate to be. When dealing with sample estimates, there is always some uncertainty about the estimate. The confidence interval is the range of that imprecision. Larger samples are needed to achieve smaller ranges. The true recall percentage will be within plus or minus  $C \times 100\%$  of the one estimated from our sample. Another way of saying this is that if we repeated the estimate with a new random sample each time, then 98% of the time, the new sample would have an estimated proportion within  $\pm C$  98% of the time. We will use 0.03 for our desired confidence interval.

The next problem is to choose a level of  $p$ , which is the proportion of responsive documents that have been retrieved. Unfortunately, we do not know this proportion before we do the analysis. In fact, it is the very thing we are trying to estimate. The worst case from an estimation point of view is when  $p = 0.5$ . As a result, statisticians often use this proportion when computing the required sample size.

$$1508 = \frac{2.33^2 \times 0.5 (1-0.5)}{.03^2}$$

Using these values, we will need at least 1,508 responsive emails to estimate recall with the accuracy we have specified. To get these we will have to review enough randomly selected documents to find 1,508 responsive ones and then count the proportion of those that were detected by the search process. This will be our estimate of the recall proportion over the whole document set. This could be a rather substantial number of documents, especially if responsive documents are rare. In any case, however, it is far fewer than all of the documents in the collection.

## XI. ELUSION AND ELUSION SAMPLING

Another, new, measure may be more easily obtained and may be more useful in the discovery context. This alternative measure also has a natural translation into a quality control process. Rather than estimating the proportion of responsive documents that have been retrieved, it may be more practical to determine whether there were significant numbers of documents that were missed by the retrieval process. This measure, called "elusion," is related, but not identical, to recall. Recall is the proportion of responsive documents that have retrieved and elusion is the proportion of nonretrieved documents that are responsive and should have been retrieved.

One can estimate elusion, but it is more valuable to use this general approach to determine whether significant numbers of responsive documents have been missed. Elusion can be used as a quality check, equivalent to the kind of quality check manufacturers would use to determine whether their manufacturing process meets their standards.

You can use standard sampling procedures to estimate the actual elusion rate, but it is usually simpler to determine whether the elusion rate exceeds a reasonable criterion. To assess elusion, you evaluate a randomly selected set of nonretrieved documents. If there are any responsive documents among the sample, you can adjust your retrieval criteria to detect these documents and then draw a new random sample. Following industrial standards we apply an “accept on zero” criterion—we only consider ourselves successful if there are no responsive documents in the sample. Elusion assesses what we missed.

The optimal sample size for this process depends on the confidence level desired and on the desired maximum probability of nonresponsive documents among the nonretrieved set—the quality standard. What is a reasonable effort expended to find responsive documents? How do we know that what we have done is reasonable?

Unlike the use of recall, we will use elusion to determine that there are no substantial numbers of responsive documents that weren’t retrieved. If there were, they were less prevalent than some specified maximum acceptable rate. To be absolutely certain that there are no responsive documents that were missed would require an infinite effort. We will have to settle, therefore, for a reasonable but rigorous level of confidence. Again, we will select a confidence level of 0.98.

The number of documents that must be sampled is determined by the formula

$$n = \frac{\log(\alpha)}{\log(1-P_s)} = \frac{\log(0.02)}{\log(1-0.02)} = 200$$

In this example, we have chosen the maximum prevalence of responsive documents in our nonretrieved set to be 2%. No more than 2% of the rejected documents are expected to be responsive ( $P_s$ ). This percentage can actually be set to any desired value. The lower the percentage, the more items have to be sampled. Reducing the maximum prevalence to 1%, for example, requires that almost 400 documents need to be sampled. The maximum prevalence you set will depend on what you think is reasonable performance of the review process. The number of documents that must be reviewed for various confidence levels and various maximum prevalence is shown in Table 1.

Table 1. The estimated number of documents to review to achieve specified levels of confidence and maximum acceptable error rates (ps).

$P_s$	Confidence			
	0.999	0.995	0.99	0.98
0.0001	69075	52981	46050	39119
0.0005	13813	10594	9209	7823
0.001	6905	5296	4603	3911
0.005	1379	1058	919	781
0.01	688	528	459	390
0.02	342	263	228	194

Based on these assumptions, 200 documents (rounded up from 194) are randomly selected from those that were not retrieved. These documents are reviewed. If any of those documents are found to be responsive, then the discovery criteria are revised to capture those responsive documents and a new sample of 200 documents is selected.<sup>12</sup> This process is repeated until the sample comes up with 0

<sup>12</sup> Technically, this repeated sampling introduces a small additional chance of error because of repeatedly sampling from the same population. This error has been ignored here, but it can be accounted for by raising the desired level of confidence.

responsive documents. Rather than merely estimating our level of success, as we would do with recall, this measure allows us to assess whether our entire process has succeeded to the level that we require. The biggest problem with using any information retrieval system is knowing what to look for. Using this elusion sampling measure allows us to assess the entire information retrieval process, including the formulation of our queries. If we have inadequately formed queries, then our elusion sample will uncover their existence and allow us to revise our criteria. There will be documents in the elusion sample and the test will have failed.

There are a number of measures of information retrieval effectiveness in addition to precision and recall. In search, these measures depend not only on the ability of the system, but on the ability of the users to derive queries that adequately cover the domain of responsive documents and on the administrators' decisions about how strict or liberal to place the criteria for retrieval. Rarely are tests conducted that would allow one to assess, from a quality perspective, the adequacy of these systems. Elusion provides one possible measure that translates directly into a quality assessment of the entire system, including the users and their queries. Elusion sampling requires only a modest amount of work that does not depend on the size of the collection, only on the specified confidence and minimum probability levels.

# THE SEDONA CONFERENCE® COMMENTARY ON EMAIL MANAGEMENT: GUIDELINES FOR THE SELECTION OF RETENTION POLICY

---

*A Project of The Sedona Conference® Working Group on  
Electronic Document Retention and Production, eMail  
Management and Archiving Special Project Team\**

Editor: Thomas Y. Allman

[REVISED April 7, 2007]

Electronic mail ("Email") is of vital importance to the productive efforts of an enterprise and its use is growing exponentially. In 2005, the average user processed 75 e-mails a day and the Radicata Group estimates that corporate e-mail traffic per user has increased at a rate of 33% per year since then. Projections are that worldwide traffic in 2006 was at the rate of 183 billion messages per day.<sup>1</sup>

Many organizations are struggling to decide how to best cope with the explosion of email while reconciling competing needs imposed by business, regulatory and litigation requirements. This Sedona Conference® Commentary suggests Guidelines for determining the core elements of an email retention policy suitable for public and private entities. Although the legal, regulatory and cultural environment of each organization varies greatly, there are common elements to a legally defensible email management policy.<sup>2</sup> In our Working Group discussions, we have been struck by the fact that entities of comparable size with similar legal risk and regulatory profiles can and do successfully adopt different retention strategies and that these strategies can vary over time, depending upon the phase of development, the size and complexity of the organization, and the particular issues most significant to the entity at any particular time.

The key is to develop and enforce in good faith those reasonable policies which best fit the entity.

The Appendix to this Commentary includes a flow chart of the decision making process and also describes two contrasting retention strategies that can be followed.

## I. SEDONA GUIDELINES ON EMAIL POLICY DEVELOPMENT

**Guideline 1: Email retention policies should reflect the input of functional and business units through a team approach and should include the entire organization including any operations outside the United States.**

---

\* Special thanks go to all the WG1 members and observers who provided valuable input and feedback. This document is for educational purposes only and is not a substitute for legal advice. The opinions expressed herein are consensus views of the editors and authors, and do not necessarily represent the views of any individual participants or authors or any of the organizations to which they belong or clients they represent, nor do they necessarily represent official views of The Sedona Conference®.

1 According to Radicati Group, the average size of email messages is also increasing, which, when combined with increased email volume, led to an increase in bandwidth storage requirements per user on the order of 61% per day from 2005 to 2006.

2 We confine our recommendations to the retention of email in its traditional sense, i.e., in its role as a means of electronic communication, often with attachments. We do not here evaluate or discuss the retention policies involved in the use of instant messaging, VoIP or other forms of electronic communication, many of which face similar issues as their usage grows.

**Guideline 2: The team should develop a current understanding of email retention policies and practices actually in use within the entity.**

**Guideline 3: An entity should select features for updates and revisions of email retention policy with the understanding that a variety of possible approaches reflecting size, complexity and policy priorities are possible.**

**Guideline 4: Any technical solutions should meet the functional requirements identified as part of policy development and should be carefully integrated into existing systems.**

## II. INTRODUCTION: FRAMEWORK FOR POLICY DEVELOPMENT

### A. General Retention Considerations

A business or public entity has the responsibility and authority to determine how best to utilize and retain information managed by its email system. This includes decisions on the specific features which impact the duration of retention of email (hereinafter collectively referred to as “email retention policies”). See *The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age* (Sedona Conference Working Group Series, September 2005)(hereinafter “Sedona Guideline No. \_\_\_\_”), *Guideline No. 3* (“An organization need not retain all electronic information ever generated or received.”).<sup>3</sup> An email retention policy may exist independently of other types of retention policies or it may be part of an umbrella policies relating to retention of information, sometimes known as “records” or “document” management policies.

Whatever their form, email retention policies must be reasonable in purpose and reasonable as applied. They must take into account any applicable statutory and regulatory mandates that may directly or indirectly govern the management of email and, regardless of those requirements, the policy must recognize the need to preserve and produce information sought in legal proceedings.

For example, the financial services industry is notable for the detailed regulations which impact the operation of information systems and mandate technical features (e.g., communications must be retained in a non-rewriteable, non-erasable format).<sup>4</sup> State, federal and local public entities are often also required to manage information created by or received on email systems in accordance with specific regulations created for that purpose.<sup>5</sup>

### B. Typical Retention Features

Email retention policies in use today frequently include some or all of the following features to maintain efficient operations. The features are not listed in any particular order based on frequency of use or efficacy in achieving their intended purpose. The focus is on management of email on active sources accessible to users<sup>6</sup> within a firewall which provides appropriate security, virus protection and reasonable spam protection.

#### 1. User Mailbox Size Limitations (“Quotas”)

Limitations or “quotas” on the amount of storage space on the network available to an individual user has historically been a principal feature of email management. In a 2005 Industry

<sup>3</sup> This principle has been endorsed by the Supreme Court in *Arthur Andersen LLP v. United States*, 544 U.S. 696, 704 (2005)(document retention policies are “common in business” and may be used to explain - and justify - the loss of information in appropriate circumstances).

<sup>4</sup> See SEC Release No. 34-37182 (May 1996). SEC Release No. 34-38245 (Jan. 1997). NASD Rule 3110, Henry L. Judd and Benjamin S. Hayes, *Records Management of E-Mail by Securities Firms: Legal and Compliance Technology Issues*. See, e.g., 17 CFR Section 240.17a-4 (originals of all communications received and copies sent by members and broker or dealers relating to its business as such must be maintained for not less than six years, the first two years in an easily accessible place).

<sup>5</sup> See, e.g., 36 CFR Section 1234.24 Subpart C. *Standards for Managing Electronic Mail Records*. Generally, Federal records are managed within a recordkeeping system and similar provisions exist at state and local levels. See Peltz, “Arkansas’s Public Record Retention Program: Finding the FOIA’s Absent Partner,” 28 U. Ark. Little Rock L. Rev. 175 (2006).

<sup>6</sup> Email which may be found on media retained for disaster recovery purposes is typically not subject to email retention policies of the type discussed herein. Any indication that backup media is routinely utilized as an unofficial archive should raise questions and call for greater analysis and review. See *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, Comment 5.b (July 2005 Version) (“Organizations that use backup tapes for archival purposes should be aware that this practice is likely to cause substantially higher costs for evidence preservation and production in connection with litigation.”).



Survey, over one half of the respondents reported that they were “managing” email retention by limiting mailbox sizes. See *Electronic Communication Policies and Procedures: A 2005 Industry Study*, AIIM & Kahn Consulting, Inc.

Implementation of quotas can vary widely. Typically, a user is given a warning before email is deleted, although the ability to send or receive email may be automatically disabled pending voluntary reduction below the size limits.

## 2. Automatic Deletion of User Mailbox Contents

Many entities provide for the automatic deletion of email after it has been retained for a certain duration (usually measured in days). Typically, an automatic deletion policy is coupled with options so that the user can move email of significance to an appropriate alternative storage location. In some instances, users are required to print and file copies of email and attachments into an (existing) records management system. Some organizations require users to assign individual retention durations to email as part of the process of classification. The process may be accomplished by “drag and drop” into a networked file with a predetermined retention duration or by making choices from drop down screens. Yet another approach is to delegate the responsibility to the user to set up and use a local file structure which mimics existing records retention schedules for that particular unit.

Automatic classification of email by content using computer generated criteria is not yet a proven technique for managing retention.<sup>7</sup>

Certain types of email content (such as that relating to contracts, regulatory files, trade secrets, and critical business records) may be sufficiently important that dedicated electronic content management repositories are also made available. *See* Cohen, “A Practical Enterprise Methodology for Addressing the Compliance Challenges of eDiscovery, eRetention Management and Defensible Disposition,” EMC Corporation (2006).

## 3. Extended Storage Options

Some entities provide extended capacity or archival storage for undeleted or unallocated email. Email not deleted by a user within a certain period may thus be saved for a long period, with decisions on moving individual email (via declaration as a “record”) to records management postponed to another day.

(It should be noted that there may be instances where an entity decides to save all inbound and outgoing email without granting the user the right to winnow out transitory or other information deemed unworthy of retention).

Entities providing extended storage typically move email from active servers to lower cost (“tiered”) storage that offer capacities such as de-duplication of messages and attachments in order to reduce server usage. A variation on this approach is to place email into a “vault” or “safe” where the user cannot change or delete the information but can continue to access it for individual use. This increases the likelihood that the information will be accessible when needed. Advocates of this approach also argue that implementation of litigation holds involving large numbers of custodians can be more easily and uniformly managed by use of this feature.

Some entities confine this strategy to specific groups – such as scientists, executives, accountants and HR representatives – because their information predictably may be needed for future business or regulatory inquiry.<sup>8</sup> For these groups, an outer limit of retention may be established based on the various regulatory and business retention requirements. These requirements stem from a large

7 *See* Frazier [Renew Data] and Brownlee, “E-Mail Management Software: The Panacea We’ve All Been Waiting For?,” Pike & Fisher Digital Discovery & e-Evidence,” (2006)(pointing out that automatic classification is also being deployed).

8 This approach is mandated in some industries, such as the financial services industry. Radicati Group reported in mid-2006 that 46% of its respondents in a corporate survey were deploying some form of archiving system and that an additional 39% of those without one were planning to deploy one in the future.

and complex series of statutes and regulations (such as tax, employee relations, payroll and benefits, environmental, etc.) as well as common-law and other business requirements.

A variety of practices exist in regard to assigning retention periods. Some entities require that users utilize existing records schedules. Others develop highly simplified versions of their records schedules (a “big bucket” approach) specifically for individual email based on content. Others assign uniform, long term retention periods to all email for classes of users or departments without individual classification on the theory that individual classification is not practicable.

#### 4. Restrictions on Local Storage

Some entities prohibit users from placing information on local PC hard drives (or other distributed devices) to avoid the problem that local hard drives cannot be accessed by others and are not backed up. Eliminating this option can be controversial, since some users may regard local storage as an essential element of their productive use of the information and may not appreciate having to rely upon networked or shared storage.

### B. The Importance of Litigation Holds

A good faith and reasonable effort to implement a “litigation hold” to preserve potentially discoverable information needed for litigation or governmental investigations must be made once a preservation obligation is triggered.<sup>9</sup>

Email retention policies which impact on retention must take this imperative into account. *See generally* Sedona Guideline 5 (“An organization’s policies and procedures must mandate the suspension of ordinary destruction practices and procedures as necessary to comply with preservation obligations related to actual or reasonably anticipated litigation, government investigation or audit”).

Courts have not hesitated to impose sanctions where email was destroyed or lost at a time when a preservation obligation was in effect.<sup>10</sup> Generally, courts have “broad discretion in choosing an appropriate sanction for spoliation,” but “the applicable sanction should be molded to serve the prophylactic, punitive, and remedial rationales underlying the spoliation doctrine.” *Silvestri*, 271 F.3d at 590 (quoting *West v. Goodyear Tire & Rubber Co.*, 167 F.3d 776, 779 (2nd Cir.1999)).

The 2006 Amendments to the Federal Rules of Civil Procedure (“the Amendments”)<sup>11</sup> do not specially spell out when or how a preservation obligation with respect to email is triggered nor do they identify the exact scope of the duty and how it is to be satisfied. Parties are instead encouraged by Rule 26(f) to assess and discuss preservation issues early and Rule 37(f) provides for limited protection from some spoliation sanctions based on the presence of “routine, good faith” operations of information systems.<sup>12</sup>

It is clear from the Rules and the emerging case law that the implementation of litigation holds involving email will receive heightened scrutiny.<sup>13</sup> Reasonable care must be taken in administering a litigation hold to ensure that routine features which may interfere with preservation are addressed.<sup>14</sup>

9 The elements of a litigation hold process in the employment litigation context were discussed in detail in *Zubulake v. UBS Warburg*, 229 F.R.D. 422 (S.D. N.Y. 2004) (“Zubulake V”). The principles applicable to a governmental regulatory action are similar to those in the civil context. *See* Cutler, Stegman & Helms, “Document Preservation and Production in Connection with Securities and Exchange Commission Investigations and Enforcement Actions,” 37th Annual Institute on Securities Regulation, 1517 PLI/Corp 579, 593 (2005) (“Retention Mechanics”).

10 For example, in *Broccoli v. Echostar*, 229 F.R.D. 506 (D. Md. 2005), a defendant that had failed to suspend an automatic deletion process upon being placed on reasonable notice of litigation was found to have engaged in spoliation.

11 The Amendments to the Federal Rules of Civil Procedure went into effect on December 1, 2006.

*See* [http://www.uscourts.gov/rules/EDiscovery\\_w\\_Notes.pdf](http://www.uscourts.gov/rules/EDiscovery_w_Notes.pdf).

12 *See* Allman, “Defining Culpability: The Search For a Limited Safe Harbor in Electronic Discovery,” 2006 Fed. Cts. L. Rev. 7, available at <http://www.fclt.org/2006fedctsr7v.htm>.

13 *See Cache La Poudre Feeds v. Land O’Lakes*, Civil Action No. 04-cv-00329-WYD-CBS, 2007 WL 684001 (March 2, 2007) (sanctioning producing party for failing to meet ongoing preservation obligations).

14 *Miller v. Holzmann*, CA No. 95-01231 (RCL/JMF), 2007 WL 172327 (D.D.C. January 17, 2007) (a litigation hold can help prevent destruction of information where a computer is programmed to destroy information after a period of time or where a person unaware of litigation may destroy electronically discoverable information).

The email retention policy may leave the specific details of litigation hold implementation to other policies and procedures, including ad hoc processes, but it is important that the entity be able to demonstrate that email retention practices are actually subject to the constraints imposed by preservation obligations as they arise. A companion Sedona Commentary to this paper is currently being prepared to address the considerations involved in establishing and implementing litigation holds.

### III. DISCUSSION

**Guideline 1: Email retention policies should reflect the input of functional and business units through a team approach and should include the entire organization including any operations outside the United States.**

**Commentary:**

Until recently, the management of email was viewed primarily as a technology issue to be handled by the IT department on advice from the legal department. However, confusion and frustration on the part of users as well as developments in the litigation context has led to some dissatisfaction with that approach.

Accordingly, it is more typical of current practice for an interdisciplinary team consisting of both functional and user representatives to be assigned to assess an email retention policy. The assessment may be driven by concerns of one of the groups, but it should include representatives of Legal, IT, Records Management (RM), Compliance, Finance, and representatives of major business units – domestic and international. Complex issues are involved when an entity operates both inside and outside the United States because its email system typically encompasses all parts of the enterprise. For example, the entity may decide to adopt decentralized email retention approaches under the umbrella of a single broad policy with individual national policies corresponding to organizational limits – or not. The team is often empowered to assess the current status of the organization's email policy and to make recommendations regarding changes, if any, that may be deemed necessary or desirable, including development of new policy features.

The team should be empowered to call upon and utilize, as appropriate, the expertise of consultants, professional organizations, outside lawyers and vendor representatives. A decision should be made as to what departments or units within the entity are going to be primarily responsible for developing, implementing and monitoring email retention policies for the functions or enterprise as may be applicable to the entity. A fully engaged responsible person should be appointed to lead the team to work closely on implementation, including recommendations on budget or funding decisions as well as monitoring the program after implementation.

**Guideline 2: The team should develop a current understanding of email retention policies and practices actually in use within the entity.**

**Commentary:**

It is important for the team to understand the types of email management policies and practices actually in place which may relate to the retention issue. The goal should be to identify the practical gaps, if any, between existing retention policies and actual practices and the costs and risks, including litigation risks, which are present. The results of this analysis can then be useful in discussions of any proposed changes or revisions under consideration.

It will be important for the team to develop an understanding of these basic background questions:

- What are the current policies, processes, work practices, or procedures applicable to the creation, distribution, retention, retrieval, and deletion of email and other electronic communications?

- What contextual information does the email system generate?
- What types of personal or distributed electronic devices are in common usage for handling email?
- What types of content are transmitted or received by email or contained within the message bodies?
- What user management practices are encouraged or tolerated for individual email accounts?
- What access to personal email archives exist on desktop and laptop hard drives and how often they are used?
- What is the role of the user in determining how long email is retained?
- When and how are existing email policies and procedures communicated to users?
- Who (and how many different functions) within the organization is responsible for or has email policies in place?
- How does the organization define a “record” and to what extent (percentage) are emails, usually based on content, included within this definition?
- How are emails with business significance, as defined by records schedules, treated?
- What are the current audit practices and capabilities to assure system integrity?
- Are users required to ascertain and classify email and to what extent is this accomplished?
- How is email integration into records management systems accomplished?
- How are litigation holds applied to email?

Some may find it useful to retain outside consulting assistance to help perform this assessment. Other entities may find it sufficient to conduct their own review of existing policies and interviews, focus groups and surveys, including benchmarking exercises or other informal methods of developing comparative examples of possible policy features.

As a starting point, IT personnel should brief the team on the relevant technology and storage architecture, including all storage locations or potential “sources” of email and attachments. Close attention should be paid to understanding the actual functioning of the active and backup email systems and the policies and practices relating to any backup media which may be utilized.

The team should also be made aware of the relevant legal principles governing the use and retention of email, with some degree of focus on the organization’s particular litigation environment. This could include a review of the types of repetitive litigation encountered, the practices followed in preserving accessible and inaccessible sources of information in contemplation of discovery in litigation, and the process followed in identifying and producing email and other information for discovery, together with any costs, burdens, and problems encountered in carrying out the discovery process.

Finally, the records management function should be asked how and whether, if at all, email and attachments are currently expected to be reviewed for content and incorporated into records management storage policies and practices. As noted earlier, the unique regulatory, business and legal requirements applicable to each entity play an important role in determining what reasonable practice is in this regard.

With this input, the team can develop a concise and accurate evaluation of the effectiveness, business needs, costs and risks associated with current practices to help set the stage for an analysis of the need, if any, for revisions in existing policies and practices.

**Guideline 3: An entity should select features for updates or revisions of email retention policy with the understanding that a variety of possible approaches reflecting size, complexity and policy priorities are possible.**

**Commentary:**

The selection of appropriate retention strategy for active email in use is the primary focus of any review. In some cases, this review may result in nothing more than establishing supplemental

processes or features to strengthen existing policy, while in others it may involve significant revisions or the development of a completely new strategy. The unique experiences of the entity within the legal, regulatory and business context will govern the types of choices available.

The process of reaching consensus in the face of differing objectives of functional and business representatives is not easy. The team members should be encouraged to openly discuss their differences of viewpoint and identify why they believe them to be important.<sup>15</sup>

Each proposed goal or objective should be assessed in the unique context in which the entity operates. The reality is that the costs and risks associated with change are often a limiting factor and the perceived benefits may not be deemed to be worth the effort and investment required. The risks associated with any particular goal should also be carefully assessed.

### **Reaching Consensus: The Default Retention Strategy**

One place to start is by developing a general consensus on the preferred duration of continued access for users (or groups of users) (the “default retention strategy”). As discussed earlier, there are sharply contrasting practices in existence today regarding the preferred approach.<sup>16</sup>

Some entities focus their strategy on quickly reducing the volume of email that is routinely accessible to users from active sources. The key mantra is that email containing only “transitory” information – of no lasting value – is to be quickly eliminated.<sup>17</sup> This approach requires that careful attention be paid to implementation of the litigation hold process. Any approach that emphasizes short retention of email should also provide alternative local or centralized storage alternatives to enable the retention of information with longer term value (“records”).<sup>18</sup>

A contrasting approach is to permit retention of email at the discretion of the user for as long as the user deems it useful (with some outer limits). Here, the argument is that an accurate assessment of the value (needed for “classification”) of individual email cannot always be made at the time of generation or receipt of email and attachments.<sup>19</sup> Some also argue that the classification decisions by individual users are, at best, rarely consistent and risk exposing the entity to loss of information needed in the case of litigation. Thus, they assert that the best way to maximize the productivity of users is to allow them to continue to have access to information, regardless of its perceived age or value.

This contrasting approach can be expensive and difficult to implement and has the substantial drawback of increasing the amount of information stored and, therefore, the amount of information which must be searched and reviewed during litigation.

### **Fine Tuning the Consensus: Choosing Features**

After discussing a basic retention strategy and the general objectives related to it, the Team can then seek a consensus on the specific features which will be included in the email policy. To help focus discussion, an attempt should be made to benchmark among similar sized organizations to see what they have found useful, taking into account the strategy they are following.

The process of identifying attractive features and assessing their risks and values should involve the full team and be undertaken in a thorough manner. Consensus will be difficult to reach because all of the viewpoints have validity and should be considered. Naturally, legal and regulatory advice will be crucial, but hopefully the full team will participate.

15 See Osterman, “E-Mail Archiving Dependent upon Corporate Culture,” Network World (March 22, 2005) (describing the resolution of competing views of legal counsel who preferred to purge and a compliance officer who preferred to save email).

16 See Ferris Research, “Email Records Management Survey: Guidelines, Technologies, and Trends” (Report 446, September 2004) (26% of respondents require quick deletion).

17 See, e.g., Kahn, “A New Approach to Records Retention,” May/June 2006, The Information Management Journal (ARMA) (suggesting that team members should [ultimately] choose between a “get rid of everything immediately” and, “keep everything forever” approach).

18 See Tottenham and Brownlee, “Records Management and Best Practices in E-Discovery,” Fulbright & Jaworski (2006)(employees drag and drop email with longer term value into networked folders and then automatic deletion can be applied to the user’s remaining inbox and sent box items).

19 See Priebe and Frankle, “Five Tenets of a Written Document Retention Policy,” Securities Litigation, Vol. 18, No. 7 (2004)(copies of most email should be saved since production from central easily-accessible electronic archive will eliminate the need to search individual PCs or backup tapes).

One approach is to ask the team to brainstorm possible features and then rank them in order of importance, taking into account the risks and costs associated with taking or not taking action. An example of how this might be done is shown below:

#### Choosing Among Alternatives

Specific Features (from a group discussion)	Costs (Tangible or Intangible)	Risks	Forced Rankings
	(Express in dollars)	(Pros and Cons)	

Once adopted, implementing an email retention policy will require careful and focused attention. Typically, the retention policy will be rolled out through targeted training and the use of internal web resources as part of compliance initiatives which will vary depending upon the culture of the entity.

As a supplement to this discussion, we include in the Appendix two alternative forms – at the extremes – of email retention strategies which give some idea of types of features which might be incorporated into a particular policy.<sup>20</sup>

#### **Guideline 4: Any technical solutions should meet the functional requirements identified as part of policy development and should be carefully integrated into existing systems.**

##### **Commentary:**

A revised email retention policy may require the purchase or licensing of additional software and/or hardware to expand storage capacity, adding new and improved search capabilities or otherwise enhancing the features of existing email systems. All these decisions involve difficult and important technical issues, including the integration with existing systems.<sup>21</sup>

There is no consensus among the Sedona Working Group members on the suitability of any particular software to effectuate retention policy, and this Commentary does not recommend any particular approach. Careful consideration should be taken to ensure that the technology is selected to meet specific goals and that policy and work practices are not being changed solely to match the available technology. The vendors who are offering such products are highly competitive and quite capable of providing targeted suggestions in response to inquires. Any assessment of the suitability of such products should be driven by business, technical and records management considerations, as discussed above, including the potential benefits to users resulting from the presence or absence of features that enhance user flexibility and access.

<sup>20</sup> None of the major professional associations, such as ACC, ARMA, AIIM or the ABA have published comprehensive collections of such policies in the private context. The most provocative source of information - thorough benchmarking - must be mined by individual teams on a case by case basis. In contrast, a wealth of examples of public policies and procedures are available for study. See, e.g., "Managing E-Mail," National Electronic Commerce Coordinating Council (2002) (Model Policy for state agencies).

<sup>21</sup> See Technical Appendix E to Sedona Guidelines (summarizing issues involving the use of electronic (digital) archives, defined as "repositories for electronic records in a form that facilitates searching, reporting, analysis, production, preservation and disposition").

Among the issues that should be examined are:

- Ability to be integrated within the entity's technology infrastructure and existing work practices
- Ability to efficiently and accurately search and retrieve information
- Technological assurance, including experience, scalability, performance and ability to integrate with existing systems
- Ability to meet user needs for productive use of electronic information
- Issues involved in migrating from the existing situation
- Costs for the hardware and software – including its acquisition, implementation and ongoing maintenance costs
- Costs associated with user training
- Professional fees needed to implement the solutions
- Ability to identify and retain information with long term records value
- Ability to demonstrate chain of custody, record integrity and respond to metadata concerns
- Ability to integrate with enterprise content management programs
- Ongoing support needs required to fully implement the policy

In the evaluation process, an entity should be aware of the differences between enterprise solutions and so-called "point solutions" – technologies focused on solving a specific problem. There are benefits and drawbacks to either approach.<sup>22</sup> Enterprise-wide solutions can be complex and expensive to implement and risk obsolescence over time.

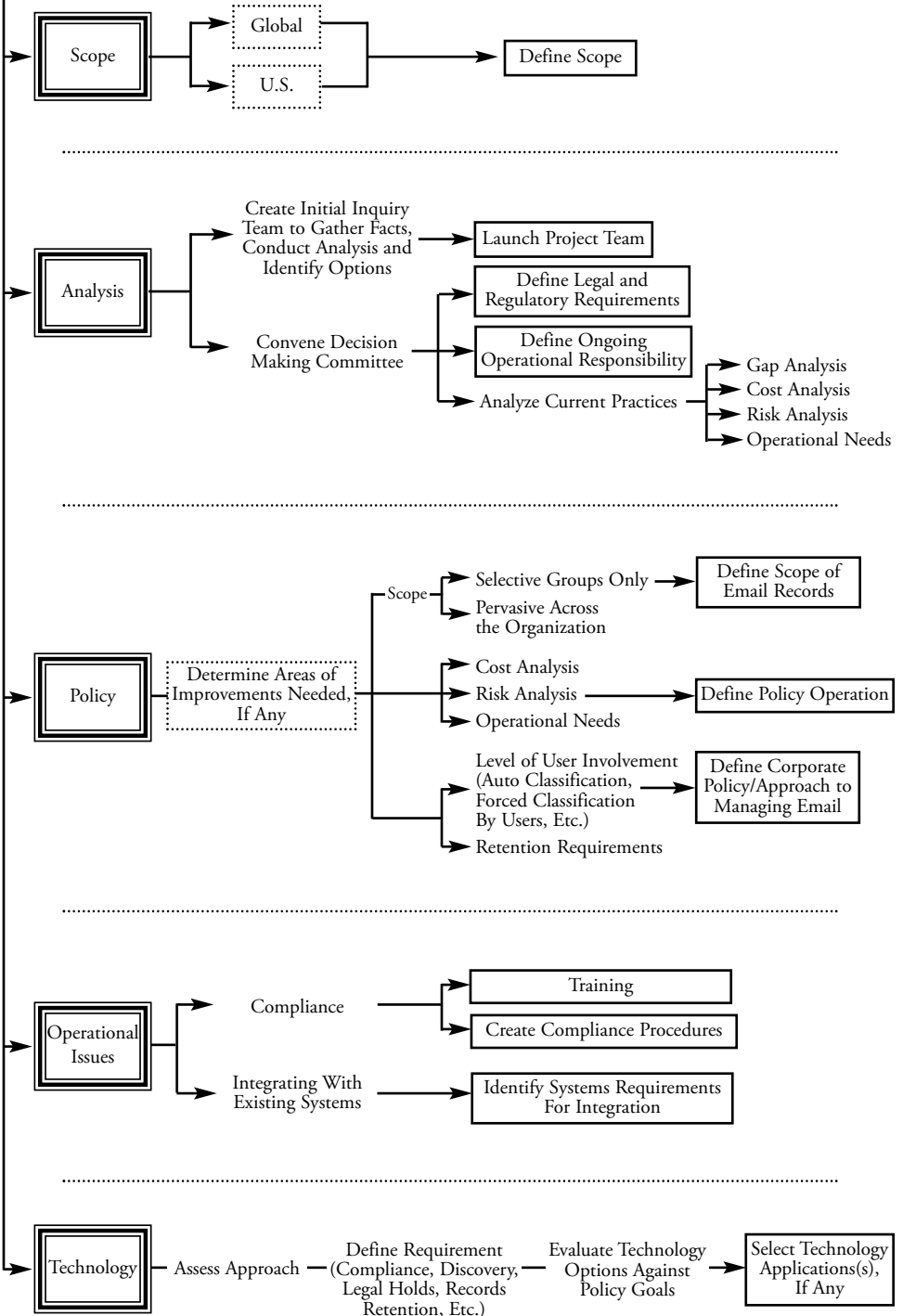
Perhaps the greatest concern is the tendency to focus on the vendor's view of what are best practices instead of focusing on implementing the entity's policy. Very few vendors have the perspective necessary to tailor their enterprise solutions to an individual entity without careful involvement of the purchaser. Similarly, while targeted solutions can be useful tools in an overall strategy, entities should only make such purchases within the context of an overall enterprise-wide strategy, as it can be difficult to integrate different solutions into existing systems.

Finally, an entity should focus on determining the true cost of adding technology to its enterprise. Some solutions may require additional development to properly customize the system to enable it to be integrated with existing disparate systems. As with any substantial purchase, performing due diligence about the vendor and requesting references can assist in making a well-informed decision.

---

<sup>22</sup> See Rugullies and Markham, "Message Archiving Becomes Part of ECM and Storage", Forrester Market Overview (2004)(message archiving decisions should be made in context of storage strategies).

### Sedona Email Management





## COMPOSITE SAMPLE POLICY FEATURES

### Policy 1.

(Based on short default retention strategy)

- **Core Policy:** Email is retained on an active server only for a short period (e.g., 30 - 90 days), which is enforced by automatic deletion and, perhaps, limits on mailbox size. The user may avoid the deletion only by taking explicit, affirmative actions such as moving the material to dedicated storage on networked files with pre-assigned retention periods. Litigation holds are applied by users to active email and the automatic deletion feature may be suspended by the entity, as appropriate, for key actors for the period until discovery is complete.
- **Related Features:** Users are expected to discard information, store it locally or on networks or print it out in hard copy and incorporate it into existing file structures subject to records management. Roles and responsibilities are clearly articulated for implementation of the policy, including management of litigation holds.

**Pros:** This approach may help reduce the rate of growth of additional primary server storage. Selected email with longer term value is retained by the user, who is best situated to understand the types of records dealt with on a daily basis and is therefore best equipped to make such classifications accurately and effectively.

**Cons:** The policy may lead to the loss of information needed by the entity in litigation (for and against the entity's position) because of the difficulty in accurately identifying the importance of particular information within such a short period of time. Users who wish to retain information for longer periods of time may find other methods (such as storing files on a local drive or the use of portable storage media) to retain valued email, thereby making subsequent review during discovery more complex.

**Legal Assessment:** While some courts are uncomfortable with automatic deletion of active email after a short period, no court has found such a process to be unreasonable where provisions for litigation holds are included and the user has alternative methods of disposition prior to deletion. If discoverable information is not preserved by a user before the copy is eliminated by automatic deletion, but after a preservation obligation has attached, a court will examine whether the use of the automatic deletion feature was "routine" and operated in "good faith," which is fact specific.<sup>23</sup>

Notably, an organization is perfectly free to choose the degree to which it relies upon the discretion of individuals in managing email and applying records schedules; It is not an indication of bad faith to rely upon individual user discretion. That said, an organization must provide those employees with adequate training and direction to exercise judgment with respect to the retention and destruction of emails.

### Policy 2:

(Based on indefinite default retention strategy)

- **Core Policy:** Email is retained on active servers for 60 days and then moved automatically to tiered storage and retained indefinitely (or a specified period such as 3 or 5 years). The user is permitted to utilize local archiving or other methods appropriate to his or her work practices. Content management and records management applications are also made available with appropriate search capability for purposes of retrieval for litigation or business use.

<sup>23</sup> See Fed. Rule Civ. P. 37(f) (effective December 1, 2006) ("Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good-faith operation of an electronic information system."). See *Turner v. Resort Condominiums International*, No. 1:03-cv-2025-DFH-WTL, 2006 WL 1990379 (S.D. Ind. July 13, 2006) ("Rule 37(f) recognizes that discovery should not prevent continued routine operation of computer systems.")

- **Related Features:** The ability to store information on local hard drives may be restricted in order to assure centralized storage of all email. A phased implementation might include archiving of legacy email (as from backup media) introduced as a second step. Roles and responsibilities, including management of litigation holds, may be specified.

**Pros:** The approach reduces the amount of data stored on the entity's email servers, increases assurance of the entity's ability to access and retrieve information for business or litigation hold purposes, and removes any motivation for users to maintain locally retained information.

**Cons:** This strategy can significantly increase the amount of stored data that must be searched and reviewed for relevancy, confidentiality, privileges and work product when subject to discovery in litigation. In addition, the costs of installation and maintenance of required systems may be large, and the complexity and reliability of some forms of archiving is still an open issue in some quarters. Finally, entities with a large amount of litigation may find it difficult to find an "open window" in its cycle of repeated litigation holds to effectively dispose of some portion of its ever growing store of email.

**Legal Assessment.** No court has held that an entity must invest in any particular form of storage technology or otherwise adopt a "save it all" preventative archiving strategy in order to comply fully with common law obligations to preserve potential evidence for pending or anticipated litigation. However, entities that chose some variation of this approach will have an argument that it is not necessary to preserve less accessible sources that may contain relevant e-mails (such as local hard drives and backup tapes) where it is most likely that such sources contain only duplicate information.<sup>24</sup>

---

<sup>24</sup> The Civil Rules Advisory Committee Note to Fed. R. Civ. P. 26(b)(2)(B) explains that "[o]ne factor [that bears on the preservation obligation with respect to inaccessible data] is whether the [responding] party reasonably believes that the information on such sources is likely to be discoverable and not available from reasonably accessible sources." *Accord: The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, Comment 5.b (July 2005 Version) ("Absent specific circumstances, preservation obligations should not extend to disaster recovery backup tapes created in the ordinary course of business. . . . [E]mploying proper preservation procedures with respect to the active system should render preservation of backup tapes on a going-forward basis redundant."); *cf. Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 ("If unique, relevant information exists on backup tapes, a party may be obligated to preserve and review such tapes.")

# THE TREC LEGAL TRACK: ORIGINS AND REFLECTIONS ON THE FIRST YEAR

---

*Jason R. Baron*<sup>1</sup>

*National Archives and Records Administration  
College Park, MD*

## *Origins*

Imagine you are involved in a multi-party case, and have just been handed an FRCP 34 Request to Produce consisting of 1,726 paragraphs, the last paragraph of which expressly says that all of the prior paragraphs – some of which ask for documents going back 50 years – apply to your client institution's records. Further imagine that you are preserving on the order of 20 million email records that would need to be searched for responsive documents to the request to produce. You cannot review every email, and must rely on automated search software to narrow the possible universe of responsive documents. How would you approach that task? Would you use a set of keywords to apply against the email database (with or without consulting your opposing party)? How do you know that the keywords you have dreamed up adequately locate all, or even a substantial percentage of, relevant email records? Where would you go to find out if there is a better, more accurate, more efficient means to conduct the search? What objective criteria would you use to evaluate how much better an alternative method might be than using keywords alone?

The scenario outlined above was not merely a hypothetical, but a real world set of concerns faced by my agency, the National Archives and Records Administration, in 2002 and 2003.<sup>2</sup> In an effort to separate truly responsive emails from false positives, NARA enlisted 25 archivists and lawyers to manually search through the approximately 1% of "hits" from the 20 million White House emails resulting from simple keyword searches. Though NARA resources were severely strained, the experience was highly instructive on several scores: first, it convinced me that simple keyword searching was a less than adequate tool when applied against very large databases (and indeed, that even a manual review of 1% of a database an order of magnitude larger would simply prove to be impossible, given present agency resources); second, were a similar tsunami wave of litigation to wash over the agency in the future, I would be recommending using more sophisticated and alternative ways of searching for evidence, including methods drawn from notions of fuzzy logic, concept searching, and statistical techniques; third, I found that there was little in the way of present-day research showing what search and information retrieval methods were objectively better to use in a legal context.

The experience of conducting a massive search for White House e-mail records in turn gave birth to the TREC Legal Track. I first conducted a literature review, and became aware of the Blair

---

1 The author serves as Director of Litigation at the National Archives and Records Administration (NARA). As NARA's representative to The Sedona Conference, the author serves as Co-Chair of the Search and Retrieval Sciences Team in Working Group 1, and as Editor-in-Chief of the *Best Practices Commentary on The Use of Search and Information Retrieval Methods in E-Discovery*. Mr. Baron is also an Adjunct Professor at the University of Maryland College of Information Studies, College Park, Maryland. The author is indebted to Ellen Voorhees and Ian Soboroff at the National Institute of Standards and Technology, and to fellow track coordinators Doug Oard and Dave Lewis, for all of their very substantial contributions toward the successful launch of the first year of the TREC Legal Track. I also wish to thank Doug Oard for his substantive comments on an earlier draft of this article, although I take full responsibility for any errors that remain uncorrected notwithstanding Doug's considerable efforts to keep me honest in fairly and accurately reporting out our scientific results. I also wish to thank Ezequiel Berdichevsky, George Washington University Law School (J.D. candidate, '08), for his expert editorial assistance. The views expressed herein are the author's alone, and do not necessarily reflect the official views of NARA, or the views of any other person or institution.

2 For a more extensive discussion of NARA's response to the discovery production in *U.S. v. Philip Morris et al.* (D.D.C.), see Jason R. Baron, "Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used In E-Discovery," 6 *Sedona Conference Journal* 237 (2005).

and Maron study.<sup>3</sup> At the same time, I became aware of the existence of TREC through discussions with NARA staff, and in late 2004 I approached the computer scientist at NIST who serves as project manager for TREC, Dr. Ellen Voorhees, to gauge NIST's interest in sponsoring research into evaluating search methods in a legal context. She was very open to the idea of a new legal track, and even encouraged NARA's last-minute participation in the then-current 2005 track year, at least on a pilot basis, if a suitable database and resources could be found.<sup>4</sup> At about the same time, I also approached my University of Maryland colleague Dr. Douglas Oard, who holds a joint appointment in the Institute for Advanced Computer Studies and the College of Information Studies, telling my tale of NARA's experience and the huge, practical problem facing lawyers. I hoped he would find the challenge interesting, and luckily he did. Later in 2005, we drew up the first draft of a proposal that would eventually be approved by Ellen Voorhees and her colleagues on the TREC steering committee.

I understood that developing a proposal for a new NIST research track devoted to e-discovery concerns in the abstract would only be the first step; executing a plan to have the track actually come to fruition meant seeking the assistance of others who would be prepared to volunteer considerable time and industry to the effort. Accordingly, I approached Richard Braman, Executive Director of The Sedona Conference<sup>®</sup>, with my ideas for a TREC Legal Track and what kind of resources I would need. He immediately recognized that this was a project that would dovetail nicely with the newly formed Search and Retrieval Sciences Team (SRS Team) within the Working Group on Best Practices for Document Retention and Production (WG1), which I had agreed to co-chair, inasmuch as the mission of the SRS Team includes assisting in future research efforts aimed at enhancing the quality of the legal process.

With the backing of The Sedona Conference<sup>®</sup>, the proposal Doug Oard and I submitted to TREC in the fall of 2005 consisted of the following three main elements: first, a sketch of possible choices for a possible test collection, including the Enron public database of email,<sup>5</sup> and the State Department cables database (*see* n.4, *supra*); second, our plan to draw on a small group of Sedona Conference<sup>®</sup> members to draft hypothetical complaints and topics; and third, our proposal for using the auspices of The Sedona Conference<sup>®</sup> to solicit lawyers, legal assistants, law clerks, law students, and other professionals with ties to the legal sector, to volunteer their time to "assess" documents for responsiveness. The proposal was accepted with the idea that we would further flesh it out at a November 2005 planning workshop, held during the 2005 annual TREC conference in Gaithersburg, Maryland.

It was David D. Lewis, Ph.D., a private consultant and long-time member of the TREC steering committee, who, prior to the planning workshop, made a strong case for us using a collection of nearly seven million publicly available documents from the "Master Settlement Agreement" database – a collection of tobacco documents produced in discovery proceedings related to lawsuits filed by the Attorneys General of several U.S. states against seven major tobacco organizations (including five tobacco companies and two research institutes).<sup>6</sup> The version of the MSA repository proposed for our use came from the Illinois Institute of Technology (IIT) Complex Document Information Processing (CDIP) 1.0 Collection, consisting of 1.5 terabytes of scanned document images, metadata, and optical character recognition (OCR) output. As described in the TREC 2006 Legal Track Overview paper ("Track Overview"),<sup>7</sup> the main strength of the collection consists of its wide range of document genres, including letters, memos, budgets, reports, agendas, minutes, plans, transcripts, scientific articles, and some email - where document length varies from

3 See David C. Blair & M.E. Maron, "An Evaluation of Retrieval Effectiveness for a Full-text Document-Retrieval System," *Communications of the ACM* 28:3 (1985), 289-299.

4 I initially presented NIST with an idea of using a database of 400,000 State Department cable telegrams from the Nixon Administration (circa 1973-74), which at the time was being prepared by the State Department and NARA for public release through a portal on the NARA web site. Although this collection represents, at best, a legacy form of test collection, *sans* email and other forms of modern-day "ESI," I was confident it would nevertheless be endlessly topic-rich and thus would be of interest for use in any long-term research project. However, a number of factors conspired against participating in the 2005 round of TREC, including technical problems NARA was experiencing with getting the database ready for public access; the cables database was, however, eventually loaded for public access in 2006, and remains a fascinating repository of government documents available to the public, as well as a possible test collection for future research.

See <http://www.archives.gov/press/press-releases/2006/nr06-71.html>.

5 The public version of the Enron collection of documents comprises over 500,000 emails with over 130,000 unique addresses; other related databases exist as well. See Tamer Elsayed and Douglas W. Oard, "The Enron and W3C Collections," powerpoint presentation at ICAIL 2007, DESI Workshop (n.13, *infra*), available at <http://www.umiacs.umd.edu/~oard/desi-ws/>.

6 For the text of the MSA and related documents, see <http://ag.ca.gov/tobacco/msa.php>.

7 See Jason R. Baron, David D. Lewis, Douglas W. Oard, "TREC 2006 Legal Track Overview," available at [http://trec.nist.gov/pubs/trec15/t15\\_proceedings.html](http://trec.nist.gov/pubs/trec15/t15_proceedings.html) (#3).

one to many thousands of pages. It was also understood from the beginning that the MSA document collection was itself the collective *output* of prior legal discovery proceedings, and thus, in some sense, is unrepresentative of a pure heterogeneous corporate collection of ESI - even one found on one of the tobacco defendant's own present-day servers. Nevertheless, as stated in the Track Overview, "[o]ur worries about that point [were] mitigated to some extent, however," by the fact that the MSA documents originated in seven different organizations in response to hundreds of document requests in multiple cases.<sup>8</sup> Speaking for myself, I perhaps was more confident than my colleagues that whatever possible bias might exist could be controlled for through the use of a range of topics as part of the evaluation exercise that, in the main, were substantially different than past MSA discovery requests.

At the November 2005 planning workshop, Doug requested that Dave Lewis and I demonstrate how lawyers might proceed to "negotiate" a search protocol consisting of a Boolean string of keywords. So far as I am aware, prior to the December 2006 amendments to the Federal Rules of Civil Procedure, it was the rare, exceptional case that involved such open, transparent negotiations over what keywords would be used by the responding party. Anecdotally speaking, I believe this to remain the case at least as of the first year after the new Rules' adoption - however, it is also my understanding that more sophisticated discussions at least amongst parties and their own consultants are also increasingly taking place as to what constitutes the proper construction of search string keywords and syntax. (Nevertheless, it remains the case, as of this writing, that there is no reported case law where parties have asked a court to intervene in connection with deciding the form of a properly-constructed Boolean string of search words to be used.) The planning workshop also established an ideal target of 50 topics for the lawyers to construct for research purposes.

### *Notes from Year 1*

One of the first tasks for the legal track consisted of constructing a usable set of topics modeled on real-world civil discovery requests. The five hypothetical complaints and associated topics in the form of requests to produce that were developed are described in more detail in the TREC 2006 Legal Overview paper. The complaints consisted of (i) an investigation by the so-called Federal Watchdog Commission into a fictional tobacco company's improper campaign contributions; (ii) a federal consumer protection lawsuit challenging a fictional tobacco company's product placement decisions with respect to television, film, and theatre shows watched by children; (iii) an insider trading securities lawsuit involving fictional tobacco executives; (iv) a state court antitrust lawsuit involving the movement of a fictional company's tobacco-related commerce in the San Diego, California area; and (v) a federal product liability lawsuit involving defective surgical devices as demonstrated in animal testing.<sup>9</sup> Care was taken to ensure that all of these complaints were readily seen as fictional in nature, even if the ultimate aim of topic development would be to find real-world documents in the MSA collection from actual companies. My thanks must go to Sedona Conference® colleagues Ryan Bilbrey, Conor Crowley, Joe Looby and Stephanie Mendelsohn, who all volunteered in drafting the hypothetical complaints and dreaming up potential topics for use in the research, and who thereafter participated in the Boolean negotiations as described below.

For each of the 43 topics used in the first year effort, I engaged in "mock negotiations" with my fellow complaint drafters, in which we essentially agreed to play the respective of opposing counsel on the propounding and receiving end of e-discovery requests. To this end, the final XML topic file that was produced as part of the track included an initially proposed Boolean query, followed by a rejoinder from the opposing party which, unless stated in the documentation, became the final agreed-upon negotiated Boolean query.

Thus, for the antitrust complaint, a typical example of a topic in the form of a request to produce and subsequent negotiations was as follows:

---

<sup>8</sup> *Id.*, Section 2.

<sup>9</sup> For a complete description of the topic set used in the TREC 2006 legal track, see <http://trec-legal.umiacs.umd.edu/> (scroll down to heading "Materials from the TREC-2006 Legal Track," subheading "Test Collection," and filename "Evaluation topics").

Topic 27: All documents discussing or relating to the placement of product logos at events held in the State of California.

Defendant counsel's initial proposal for a Boolean string of search terms for this topic was: ("product placement" AND logos AND California). Plaintiff counsel countered with [{"product placement" OR advertis! OR market! OR promot! AND (logo! OR symbol OR mascot OR marque OR mark) AND (California OR cal. OR calif. OR CA)}].

Clearly, many other terms could have been substituted into the proposed string. However, one must start somewhere, and for purposes of the TREC Legal Track the outcome of these negotiations over what constituted one form of a reasonable Boolean string was intended to serve the purpose of providing a baseline, both for track coordinators as well as the participants in the track, as to how one might approach the more general search problem posed by each topic query. For the track coordinators, it was essential that someone perform a "Boolean run," to obtain what were deemed "baseline" results from the MSA document collection using the final negotiated Boolean string; theoretically, each participant in the track would then be free to utilize and build upon the expressly stated Boolean negotiated query, in performing their own alternative ways of searching through the test collection for potentially responsive documents.

Thereafter, the success of the first year of the Legal Track was dependent on (a) attracting a nucleus of participating institutions to submit "runs" based on whatever home-grown search methodologies they wished to use; (b) pooling the results of the submitted runs so as to create properly sized "judgment pools" for the assessment phase of the track; and (c) engaging a sufficient number of volunteer assessors to assist in reviewing actual documents. Six institutions from around the world (Hummingbird, now Open Text; the National University of Singapore; Sabir Research; the University of Maryland; the University of Missouri-Kansas City; and York University), submitted a total of 31 ranked runs with no more than 5,000 documents per topic. In addition, a run known as the "expert manual searcher run" was also employed, consisting of a research assistant performing her own searches of the test collection through automated means, leading to production of up to 100 documents per topic from the collection that she felt would be unlikely to be retrieved by fully automated systems. Computer scientists at NIST, working with Dave Lewis, performed the necessary statistical stratification to come up with combined judgment pools per topic; these consisted of on average about 800 documents per topic, collected across all of the submitted runs.

The assessment phase of the track began in earnest during the first week of August 2006, in which 35 volunteers, including participants from Sedona Conference® member law firms and legal technology companies, NARA staff, and participating law schools, assessed a total of 32,738 documents for 40 completed topics. The volunteers included eight lawyers, ten law students, three paralegals, one professional archivist, one historian, and several individuals with degrees in science or finance. The affiliations of volunteers included Sedona Conference® members Bank of America, FTI Consulting, H5 Technologies Inc., Lewis & Roca LLP, Preston Gates LLP, as well as additional volunteers from NARA, George Washington University Law School, George Mason University School of Law, Reasonable Discovery LLC, the New Mexico State Attorney General's Office, and three private individuals. The Overview Paper provides a more complete summary of the problems and issues that arose in the assessment phase, which ran fairly smoothly over the course of approximately six weeks. Individual assessors were given a "how to" guide for working with an online judging platform, and individual questions were responded to by track coordinators on an ad hoc basis. Although some effort was put into defining relevancy, assessors were basically instructed that a document should be considered relevant when the reference to a topic was found in the document, and to use their best judgment. A special rule that required leaving the document unjudged if relevance could not be easily determined upon cursory examination was imposed for the occasional document that surpassed 300 pages.

A passage from the Overview paper describes the experience of assessors in somewhat more detail:

Some of the assessors went beyond the text of the topic (the complaint, the production request, and the Boolean queries) to perform additional legal research which they viewed as helpful to the exercise. For example, the assessor for Topic 30 researched at greater length what the numbered statutory code provisions were corresponding to the California Cartwright Act, to ensure that all documents containing such references, with or without reference to the Cartwright Act itself, would be marked as responsive. The assessor on Topic 10 performed independent research into the ban on tobacco advertising, as an aid to understanding what documents might be expected to be found in response to a topic involving tobacco product placement in television or film. One assessor asked for assistance on the definition of one of the keywords in the topic, leading to additional research conducted on the Internet.

Some differences were observed in how liberally or narrowly assessors viewed the scope of their discretion to find responsiveness. In some exceptional cases, assessors were willing to find responsiveness even where a key term might be missing, if the document was otherwise sufficiently generic and might yet be viewed as responsive with the aid of further research. For example, the assessor for Topic 9 (“All documents discussing, referencing or relating to payment of compensation to 20th Century Fox Corporation for placement of products and/or brands in a film production”), marked certain documents as relevant even if the film company was not expressly mentioned, where the context indicated that the company might be involved. In most cases, however, assessors appeared to adopt relatively restrictive interpretations on what met the mark for relevance.

Assessors reported some confusion as to whether they should exclude documents that might be within the literal scope of a production request when read in isolation, but which weren’t relevant to the main thrust of the associated complaint (i.e., the document had nothing to do with the causes of action in the lawsuit or investigation). The question of scope arose in particular for production requests associated with the one complaint that on its face did not involve allegations against the tobacco industry (but which was instead about medical devices). Topic 49, which coupled that complaint with a production request for “[a]ll documents created between 1962 and 1999 referencing or including warnings or draft warnings used in the United States,” proved to be particularly problematic because it was read by the assessor as being aimed at warnings for faulty medical devices. Not surprisingly, no relevant documents were found for topic 49. It was therefore removed from the evaluation because topics with no known relevant documents can not be used to compare the effectiveness of alternative system designs. Results are therefore reported for the remaining 39 topics.

As is often the case, assessors found some unintended ambiguity in the topics, either due to grammatical construction of the topic (e.g., what did the word “their” refer to), or due to inherent ambiguity embedded within words or concepts (e.g., what constitutes “lobbying efforts,” “advertising,” “marketing,” and “promotion”). For one assessor, the word “event” (in a topic asking for all documents relating to the placement of product logos at events held in California), prompted them to consult the Random House Dictionary, where the word is defined as “something that occurs in a certain place during a particular interval of time.” Therefore, in this assessor’s view, documents that mentioned such activities as the America’s Cup Race, speed skiing, auto racing, Hispanic Cultural events, Swing jam weekend, an antiviolence campaign, a country music festival, and an anti-smoking campaign called “Tobacco is Whacko,” were all properly within the scope of the topic.

Another miscellaneous concern of one or more assessors involved how to deal with documents containing foreign language text. The track coordinators instructed assessors to make judgments based on English portions of documents, or otherwise mark the document as unsure.

In general, assessors took their jobs very seriously. A number of assessors made a second pass through their document set to resolve anomalies or to revisit judgments based on knowledge gained on the first pass.<sup>10</sup>

Recruitment for assessors remained a substantial concern through the months of July, August and September 2006, especially towards the end when my fellow track coordinators made the surprise announcement that we really should also be performing a “dual assessment” round, meaning that additional volunteers would be needed to review a subset of the documents already assessed by earlier volunteers. However, my anecdotal experience from the first year confirms that lawyers and law students are actually more than willing to perform otherwise tedious document review if they believe - correctly, in this case - that they are advancing legitimate scientific research. They even reportedly (based on post-assessment surveys) have some amount of fun in doing so. I certainly have been in no position to correct them or suggest otherwise!<sup>11</sup> I trust that the experience bears out for the second year of the track, where we are initially targeting law students to volunteer in performing assessment duties for a new round of 50 new topics based on four new hypothetical complaints.

### *Scientific Findings*

The first year of the TREC Legal Track has been the subject of the aforementioned Overview paper, four additional written submissions by track participants to the 15th TREC Proceedings,<sup>12</sup> two research papers presented at a recent conference on search and retrieval issues in Palo Alto,<sup>13</sup> and at least two other Ph.D. students are known to be currently working on additional scholarship. Subject to an important caveat, I will here venture to summarize some of the findings from the first year. The caveat: the Overview paper the track coordinators initially produced does not purport to represent a comprehensive analysis of the first year data. Rather, at best, it constitutes a preliminary stab at evaluating the first year's results after performing some data analysis. Accordingly, the present piece should not be viewed as anything near a comprehensive report on the findings of the TREC Legal Track. Much more data analysis and revisiting of what actually has taken place in year 1, possibly combined with the results of year 2 and future years, will be needed, before more definitive conclusions emerge from the results of the overall study. That said, my summary follows:

First, none of the participating institutions submitting automated runs based on alternative ways of performing searches was found to have “the” answer to the search problem faced by lawyers. To the contrary, the results accompanying Figure 3 of the Overview paper show that the best runs from three of the participating institutions were nearly indistinguishable under one widely-reported statistical measure (mean R-precision),<sup>14</sup> and one of those runs (submitted by Hummingbird), was artificially limited in its role as the *baseline* Boolean run. Indeed, it might well have been surprising if one automated method was better at performing searches in a statistically significant way than all others. The extent to which no individual run did better than the baseline Boolean run is, however, somewhat curious, counter-intuitive, and deserves of further reflection.<sup>15</sup> The Overview paper makes

<sup>10</sup> TREC 2006 Legal Track Overview, *supra*, n.7, at Section 4.2.

<sup>11</sup> The experience may be perhaps likened to the “reward” of performing endless Bluebook citation checking, a task more-or-less enthusiastically embraced by generations of newly recruited members of law review.

<sup>12</sup> See generally TREC Legal Track web page for research papers, available at <http://trec-legal.umiacs.umd.edu/> (under “2006 Results”).

<sup>13</sup> See written submissions to the Eleventh International Conference on Artificial Intelligence and Law, Supporting Search and Sensemaking for Electronically Stored Information in Discovery Proceedings (ICAIL 2007, DESI Workshop), Palo Alto, CA, June 4, 2007, available at <http://www.umiacs.umd.edu/~oard/desi-ws/> (under “Research Papers”). See also Jason R. Baron and Paul Thompson, “The Search Problem Posed By Large Heterogeneous Data Sets in Litigation: Possible Future Approaches to Research,” ICAIL 2007 Conference paper, available at <http://www.umiacs.umd.edu/~oard/desi-ws/> (under “Related Paper presented at Main Conference”).

<sup>14</sup> “Precision” is the measure of the relevant documents that have been retrieved, as a percentage of the total number of documents that have been retrieved, at any given point or “rank” in a search. As defined in the Overview paper, “mean R-precision” is computed as an average across all topics of the density of relevant documents at rank R, where R is the number of known relevant documents for each topic. TREC Legal Track Overview, Section 5.2 (“R-Precision”).

<sup>15</sup> For further discussion of this point and an excellent overall technical analysis of year 1 of the Legal Track, see Stephen Tomlinson, “Experiments with the Negotiated Boolean Queries of the TREC 2006 Legal Discovery Track,” [http://trec.nist.gov/pubs/trec15/t15\\_proceedings.html](http://trec.nist.gov/pubs/trec15/t15_proceedings.html) (#57).



the point that with only six participating institutions, “we are nowhere near exhausting the design space for search techniques, so ways may yet be found to achieve improvements that are not available to a Boolean system.”<sup>16</sup>

A second result from the same data shows that the “expert manual searcher” outperformed all of the automated runs under a mean-R precision measure. While an important finding, one must be careful not to over-interpret this result: the manual search being referred to was one research assistant’s interactive attempts to construct 100 documents per topic, using an automated interface for conducting searches, and based on her superior knowledge of the MSA collection gathered from prior research. The outcome of her efforts was emphatically *not* the same as if someone with no *a priori* knowledge of the collection performed a straight “manual” review of some subset of documents. As lawyers, we must be very careful to understand how the term “expert manual searcher” has been deployed.

A third, most intriguing result, emerges in the Overview paper’s section entitled “Uniques analysis.” As the paper relates, “[o]ne way of characterizing the results of different approaches to searching is to examine the contribution of each approach to the total set of known relevant documents.”<sup>17</sup> Figure 2 of the Overview paper proceeds to show that “on average across the 39 topics, 57% of the known relevant documents were found by the reference Boolean query (i.e., either uniquely by the reference Boolean system, or by the reference Boolean system and also one or more other systems).”<sup>18</sup> Another 11% of relevant documents were found by the expert manual searcher. This leaves 32% as the combined average amount of known relevant documents across all 39 topics that were found by systems other than the reference Boolean system.

What are we to make of this figure, that I have elsewhere called “dark matter” potentially representing the relevant documents not found by simple keyword searches or other traditional means?<sup>19</sup> Coming out of Year 1 of TREC, the figure represents a combined result - a pooling of *all* of the other search methods employed by Year 1 participants beyond the “baseline” Boolean method - and thus is useful only when understood in that important context. That is, the figure does not show that any particular alternative method will guarantee that a relevant document can be found. On the other hand, if anything, the figure represents a lower bound on the number of undiscovered relevant documents that might possibly exist were still other means employed to search for them.

My fellow Legal Track coordinators would caution that any strong claims regarding what this data represents are at best premature, based on the level of analysis conducted to date. Nevertheless, I wish to suggest that the legal community should sit up and take notice of what this set of empirical data from Year 1 may yet prove to show. Lawyers historically have accepted results obtained through the use of simple searches, based on keywords alone. More recent scholarship has pointed the way towards showing that despite the power of keywords, there are known deficiencies in over-reliance on simple keyword searches alone as the method of choice.<sup>20</sup> The 32% figure represents a universe of potentially relevant evidence not found in a concededly large haystack of ESI - even after parties had engaged in a relatively sophisticated set of negotiations with the outcome of the reference Boolean string of terms. This suggests, at a minimum, that parties should be open to experimenting with alternative forms of search methods and techniques, and using them in a combined fashion, so as to maximize the rate of recall to be obtained, *i.e.*, maximizing the proportion of all relevant documents found as the result of any particular search or searches.<sup>21</sup>

16 TREC Legal Track Overview, Section 5.2. The paper goes on to describe the fact that not all Boolean runs are created equal, which reflects the fact that, in one case, the Boolean run used the initial rather than the final queries, and in other run there was incomplete or inconsistent automated support for implementing extended or non-Boolean operators (such as where the topic requests used truncated terms, *i.e.*, “!” to denote all possible further suffixes to a particular keyword or portion thereof, and/or where the topics employed the device “w/3,” *i.e.*, “within 3”). *Id.*

17 *Id.*, Section 5.1.

18 *Id.*

19 See Jason R. Baron, “Thinking Outside The Boolean Box: Metastrategies for Conducting Legally Defensible Searches In An Expanding ESI Universe,” powerpoint presentation at ICAIL 2007, DESI Workshop, available at <http://www.umiacs.umd.edu/~oard/desi-ws/>.

20 See “The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery,” Section 3, 8 *Sedona Conf. J.* -- (2007) (“Sedona Search Commentary”); George L. Paul and Jason R. Baron, “Information Inflation: Can the Legal System Adapt?,” 13 *Rich. J. Law & Tech.* 10 (2007), at Paragraph 37, available at <http://law.richmond.edu/jolt/v13i3/article10.pdf>.

21 See *Sedona Search Commentary*, *supra*, n.20, at Section 4; G. Paul & J. Baron, *supra*, n.20, at Paragraph 41.

Much greater exploration of this form of “dark matter” is obviously necessary, particularly in terms of what kinds of searches are most likely to produce unique relevant documents not otherwise found by traditional keyword searches and/or Boolean techniques. In doing so, we may take a significant step towards performing the type of objective benchmarking of competing search technologies that I have elsewhere suggested is imperative, given a fundamental desire on the part of one or both parties in litigation to understand how well e-discovery obligations are being met, as well as in providing assistance to the trier of fact to determine what has constituted a “reasonable” search for responsive evidence.<sup>22</sup>

The TREC Legal Track also has produced empirical data on two “process”-oriented issues that may be of general interest. First, some attempt was made to assess the effects of differing assessor interpretations, through the use of what is known as “dual assessments.” Based on a small sample of 50 documents per topic, the study showed a moderate rate of overall agreement between assessors (where so-called “kappa values” can range from -1 for complete disagreement to +1 for complete agreement, the mean value of kappa for the legal track was +0.49). In the case of Topic 24 (“All documents referencing the Federal Election Commission dated subsequent to 2001”), there was actually a slight negative kappa figure (-0.037), indicating more disagreement than agreement as between two assessors on their respective relevance judgments. The entire matter of the divergence of views among law students, paralegals, junior, and senior lawyers in judging relevant documents for relevance is a subject wide open for future research.

Finally, another form of “process” metric was tabulated from post-assessment survey reports, namely, how much time assessors spent in performing their document reviews. (Recall that documents could range from 1 to over 300 pages, with the 300 page figure representing a cut off for performing the normal expected full-text review.) Based on time data from 16 of the 35 participants, representing 39% of the overall assessment effort (12,743 out of the 32,738 assessments), the reported review rate per hour ranged from a low of 12.33, to a high of 67.5, with an average of 24.7 documents per hour. This figure certainly has great value for estimating the commitment levels volunteers will need to make in future iterations of the legal track, but may also be of interest to a greater community of researchers.

The data gathered during the first year of the track will be subject to continued study for years to come, both on its own and in conjunction with data gathered in the second and any future years of the existence of the track. I welcome all such efforts. Indeed, the value of the TREC Legal Track exercise is that the test collection, protocols, fundamental methodology, and research results are all transparent and openly available to the world to replicate (and improve upon).

### ***Additional Observations***

As my computer science colleagues informed me near the end of the first year’s cycle of work, the first year of a new TREC track is intended to function as an experimental pilot, where the bugs and kinks are worked out, in order to perform solid research in second and succeeding years. (Had I only known!) Although we were able to push out a 2006 Legal Track Overview paper, as stated above, substantive scholarship analyzing the results of the first year of the track, combined with the results obtained in the second and any succeeding years, largely remains to be undertaken.

As the Overview paper notes, “[p]erhaps the greatest accomplishment of the TREC 2006 Legal Track is that it happened at all. More than 50 volunteers contributed to assembling and distributing the collection, creating topics, developing systems, managing submissions, creating pools, judging relevance, developing metrics, creating scoring software, analyzing results, and coordinating those activities. This has yielded the results that we would hope for from any TREC track in its first year: (1) a reusable test collection to support future research, (2) a set of baseline results to which future research can be compared, and (3) a community of researchers who bring a variety of perspectives to these important challenges.”<sup>23</sup> The fact that the Legal Track exists, and that this type of

<sup>22</sup> See J. Baron, *Toward A Federal Benchmarking Standard*, *supra*, n.2. G. Paul and J. Baron, *Information Inflation*, *supra*, n.20, at Paragraph 46.  
<sup>23</sup> TREC Legal Track Overview, *supra*, n.7, at Section 6.

study holds the potential to aid in objectively modeling present practices in the area of civil e-discovery, continues to be enormously professionally rewarding.

I do have a continuing concern about the methodology of pooled relevance assessments employed as part of TREC. As stated in my paper with Paul Thompson,

The assumption was made that if a large number of systems participated in TREC that by judging for relevance the top n documents retrieved by each system this pooled set of judged documents would provide a representative sample of the document collection as a whole, particularly if many systems with diverse ranking algorithms participated. This assumption seemed reasonable, but is now breaking down as collections are becoming increasingly large. The problem is made worse if there are also a large number of relevant documents for each query. For example, assume that a document collection has one billion documents and that around one million of these may be relevant for a given query. Judging the top several hundred documents retrieved for each query for each of 20 or 30 participating systems will not give a good estimate of recall for the collection as a whole.<sup>24</sup>

In other words, I continue to hope that, either as part of TREC or in additional research to come, the real-world concerns of lawyers facing the need to find responsive ESI in exponentially greater universes of data will continue to be adequately modeled. Nevertheless, as I have made clear, I believe the TREC legal track project itself holds the potential to produce useful and important results of benefit to the legal profession as a whole, as we all struggle to make sense of litigation demands and obligations in this brave new world of ESI.

---

<sup>24</sup> See J. Baron and P. Thompson, *The Search Problem Posed By Large Heterogeneous Data Sets in Litigation*, *supra*, n.13, at Section D.

# NOT YOUR MOTHER'S RULE 26(F) CONFERENCE ANYMORE

---

*Moze Cowper and John Rosenthal*<sup>1</sup>

## I. INTRODUCTION

On December 1, 2006, the way we conduct federal litigation (and in many instances state litigation) dramatically changed in that we are now required to preserve, collect and produce all forms of electronic documents (*e.g.*, e-mail, MS Word, MS PowerPoint, databases), or as referred to under the new Federal Rules, "electronically stored information" or "ESI." For the first time, all parties will be under an affirmative obligation to understand and litigate technical issues concerning ESI.

Perhaps no change in the Federal Rules is more dramatic than the one associated with the changes to the meet and confer requirement of Rule 26(f). Litigants and their counsel that fail to heed the significance to the changes in Rule 26(f) will not only be doing their clients a disservice, they may dramatically increase the litigation costs and potential risks faced by their clients regarding the discovery of ESI.

This article addresses the implications of the changes to Rule 26(f) and provides you guidance regarding best practices to prepare for and conduct the Rule 26(f) Conference. We attempt to do so by taking into account two different perspectives: (i) the inside counsel responsible for the corporation's preservation, collection and production of electronic records; and (ii) an outside counsel responsible for implementing e-discovery response plans, including the preparation for and conduct of the Rule 26(f) Conference.

## II. CHANGES TO FED. R. CIV. P. 26(F) AND THEIR IMPLICATIONS

One of the primary changes to the Amended Federal Rules are to require parties to focus upon and discuss the discovery of ESI early in the litigation process through Amended Rules 16(b) (regarding the discovery planning conference with the court), 26(a) (regarding the initial disclosures) and 26(f) (regarding the obligation to meet and confer in developing, among other things, a discovery plan). The intended goal of these amendments is to encourage, if not mandate, parties to work together in addressing the difficult and complex issues regarding the preservation, collection, review and production of ESI.

Federal Rule 26(f) requires parties to confer "as soon as practicable," but in any event at least 21 days before a scheduling conference is held or a Federal Rule 16(b) scheduling order. There are three primary changes in the Amended Rule. First, the parties are required to discuss the issue of preservation of potentially relevant records. This obligation applies not only to ESI but also to all "documents" as that term is defined within the Federal Rules. This is the first time the obligation to discuss and address preservation has ever been expressly set forth in the Federal Rules. Second, the

---

<sup>1</sup> Moze Cowper is Senior Counsel at Amgen Inc. John Rosenthal is a partner with the law firm of Howrey LLP. Messrs. Cowper and Rosenthal have extensive experience in the practical application of electronic discovery. Both are participants in The Sedona Conference Working Group One on Electronic Document Retention and Production. See [www.thesedonaconference.org](http://www.thesedonaconference.org).

parties are required to discuss the disclosure of ESI and, specifically, the format production of ESI. Finally, the parties are required to discuss the issue of privilege in the context of inadvertently produced documents.<sup>2</sup>

### III. IMPLICATIONS OF THE CHANGES

Historically, Federal Rule 26 Conferences have been *pro forma*, accomplishing little if anything of significance in terms of the conduct of the case. Essentially, counsel for each side showed up at the meet and confer, which everyone knew in advance would take about fifteen to twenty minutes. In the meeting, the parties agreed to a general discovery schedule, but agreed to disagree on everything else relating to discovery.

Why were these meet and confers of little significance? The answer is very simple: defendants (typically larger corporations) had little to gain in the Rule 26(f) Conference. In other words, the meet and confer was usually an opportunity for the plaintiff to engage in unfettered discovery with little upside for the defendant corporation. The result is what one federal judge refers to as the “drive-by” discovery conference, where little if anything was accomplished.

Most commentators agree that the changes to Rule 26(f) will dramatically transform the initial meet and confer process from a meaningless *pro forma* meeting in which little or nothing is achieved to a substantive meeting whose results will directly impact the conduct of the entire litigation. There are several reasons behind this dramatic transformation.

First, defendants (particularly large corporations) now have something to gain from the meet and confer by negotiating down the scope of the discovery of ESI and, thus, the associated risks and costs. For example, the conference can now be used to limit the time period, the systems and the number of custodians at issue. In a larger case, this can result in savings in the tens of thousands.

Second, plaintiffs also have something to gain. It is expensive not only to produce ESI but also to review ESI. The more sophisticated plaintiffs understand that if they are unreasonable in the scope of their request, they may just get their worst nightmare – what they asked for. Plaintiffs, therefore, have an incentive to negotiate down a reasonable scope of the discovery of ESI that will, in turn, limit the costs associated with the document review. This is particularly true where both the plaintiff and defendant are corporations and, thus, unreasonable and overbroad discovery requests they issue are likely to be put into the word process and turned back onto them.

Third, the Rule 26(f) Conference is a means to frame the issues for the Rule 16 Discovery Planning. It is important to set out a reasonable position in the face of an unreasonable requesting party in order that the Rule 16 Conference can be used to obtain some limitations on the scope and amount of the ESI to be discovered.

### IV. PREPARATION FOR THE RULE 26(F) CONFERENCE

It is axiomatic that prior planning prevents poor performance. This is true both in life and in the context of the Rule 26(f) Conference. If the goal is to have a productive Rule 26(f) Conference that will address preservation, scope of production, production format and privilege, both inside and outside counsel need to take the time to educate themselves regarding: (i) the information technology systems at issue; (ii) where and on what systems the ESI is located; (iii) who are the

<sup>2</sup> Federal Rule 26(f) provides:

The parties must... confer to consider the nature and basis of their claims and defenses and the possibilities for a prompt settlement or resolution of the case, to make or arrange for the disclosures required by Federal Rule 26(a)(1), to discuss any issues relating to preserving discoverable information, and to develop a proposed discovery plan that indicates the parties' views and proposals concerning...

(3) any issues relating to disclosure of electronically stored information, including the form or forms in which it should be produced;

(4) any issues relating to claims of privilege or of protection as trial-preparation material, including - if the parties agree on a procedure to assert such claims after production - whether to ask the court to include their agreement in an order...

Fed. R. Civ. P. 26(f).

custodians/owners of the data; (iv) the steps undertaken by the party to preserve the relevant data; and (v) the scope of production that the party is willing to undertake (*i.e.* will search terms or date restrictions be used to narrow the universe of documents?); (vi) the ultimate form of production (TIFF, PDF, *et cetera*); and (vii) privacy considerations (*i.e.*, is the data located outside of the U.S.?).

How do you obtain this knowledge? In the perfect world, the party has undertaken steps before this litigation to understand and, at some level, to inventory its computing systems and the ESI maintained on those systems. Regardless of the existence of such inventory, inside and outside counsel are going to have to work with the information technology, records management and/or compliance departments (collectively the “e-discovery” or “litigation response” team) to obtain an understanding of these issues as a general matter and, significantly, in the context of the particular litigation. We recommend that in larger- to medium-sized cases that the e-discovery team meets in-person in advance of the Rule 26(f) Conference to discuss these very issues and begin the process of preparing for the meet and confer. We find that this meeting pays for itself a hundred times over in terms of the ultimate cost savings that can be obtained through using the Rule 26(f) Conference to limit the scope of the case.

## V. GROUND RULES FOR THE RULE 26(F) CONFERENCE

The success of the Rule 26(f) Conference will often depend upon certain basic ground rules. Here are ours:

1. To get something, you have to be willing to give up something – in this case, information about your ESI. In other words, if the Rule 26(f) Conference is to be meaningful, the parties must be prepared to engage in a good faith exchange of certain basic information about their respective ESI. This may also include exchanging detailed information about your custodians (*e.g.*, organizational charts) or computing systems.
2. People play much nicer when you are “off-the-record.” The parties will have to decide whether their Rule 26(f) Conference discussions will be on- or off-the-record. Our experience is that these conferences are far more productive when the parties feel they can engage in conversations and exchange information without the fear they will be quoted back to the court. Regardless of which way you choose to go, it is important to have a clear understanding with your opponent prior to the Rule 26(f) Conference.
3. Lawyers may not always be the most important participants. Many lawyers do not have the requisite technical understanding of either e-discovery or the computing systems at issue in the litigation. In some instances, a member information technology department or an outside e-discovery consultant may be better equipped to understand and address the preservation, collection and production of ESI than the trial attorney. In other instances, it may make sense to have a national e-discovery coordinating counsel that can attend or manage the Rule 26(f) Conference. A strategic decision will have to be made as to who, beyond trial counsel, should attend and whether they will attend to provide advice to the trial attorney or as active participants in the Conference.
4. Set out the expectations in advance. The Rule 26(f) Conference will not be productive unless you set out the ground rules and topics for discussion in advance. Send the opposing counsel a letter outlining: (i) the ground rules; (ii) the topics to be addressed and possibly your preliminary position on certain topics; (iii) what materials the parties should be prepared to bring and exchange (*e.g.*, organizational charts); and (iv) who will be in attendance and their anticipated roles.
5. Days of the “drive-by” Rule 26 Conference are over. If the Rule 26(f) Conference is to be successful, it will probably take more than one meeting. The parties will have to meet and confer on several different occasions in order to reach meaningful

agreements on preservation, scope of production, production format and privilege. Our recent experiences indicate that in larger cases, these meet and confers can stretch out over a month or two. This, of course, assumes that both parties are willing to work in good faith to achieve a reasonable compromise on the scope of the e-discovery to be conducted.

6. To adapt a famous quote, “keep your friends close, and any document agreements with your adversary closer.” The purpose of the Rule 26(f) Conference is to ultimately reach agreements that will reduce the costs of litigation while forcing the parties to pay attention to what they ultimately need to prove or disprove. To this end, it is important that the parties memorialize the agreements resulting from the Rule 26(f) Conference in order that there be no confusion about each side's obligations as they move forward in the discovery process. This can be done either formally by the Proposed Rule 16(b) Order or, alternatively, through a stipulation and agreement.

## VI. RULE 26(F) CONFERENCE DISCUSSION TOPICS AND STRATEGIES

Of course, any topic is fair game for the Rule 26(f) Conference, but the following are the ones we feel are the most significant to be addressed as either: (i) mandatory topics under the Rules; and/or (ii) practical consideration in reaching a reasonable agreement on the scope of e-discovery.

### 1. Preservation

As noted above, preservation of both paper records and ESI is a required subject of the meet and confer. More importantly, you can use the Rule 26(f) Conference as a means to limit the potential exposure to future spoliation claims by reaching an agreement on the scope of the preservation to be undertaken and, specifically, how to limit the scope of the preservation obligations through limitations on: (a) date ranges (both front-end and back-end cut off); (b) specific custodians; (c) specific systems; (d) specific categories or types of ESI; and (e) search terms. Presumably, these very same limitations would ultimately apply to the collection and production of ESI, although the scope of preservation is often broader than the scope of production.

#### a. Date Range

Opposing counsel often requests information from X date to present. Without a back-end discovery cut-off date, this places a great burden on a corporation because its employees must preserve documents as they are created, sent or received during the pendency of the litigation. By obtaining front-end and back-end cut-offs, that burden is substantially reduced because the company then has a defined date range of materials that only need to be preserved and collected once during the litigation. Accordingly, the Federal Rule 26 Conference should be used to secure both a front-end and back-end discovery cut-off for both the preservation and production of ESI. Make it clear that documents before and after that date will not be preserved. If opposing counsel is unable or unwilling to agree to both front-end and back-end cut-offs, then attempt to narrow the scope of those documents for which an open-ended back-end cut-off will apply.

#### b. Custodians

Custodians are the people who have custody of the potentially relevant documents. During the 26(f) Conference, the parties should attempt to agree on particular custodians for which documents will be preserved and ultimately produced. Limiting preservation to particular custodians will substantially ease the burden of e-discovery. You will also find that most courts think that parties engage in too much discovery and, therefore, the courts are likely to impose a reasonable limitation on the custodians at issue to the extent the parties are unable to achieve a reasonable compromise through the Rule 26(f) meet and confer.

### c. Systems

“Systems” refers to the electronic system where electronic data is stored. Possible systems include, but are not limited to: (1) e-mail systems; (2) core office document systems (MS Word, Lotus Notes, MS PowerPoint); and (3) application and databases (*see* further discussions of databases and application *infra*). During the 26(f) Conference, counsel should try to identify particular systems that will be subject to discovery versus those that will be exempt from discovery. Again, it may be necessary to engage in informal discovery or informal exchanges with opposing counsel regarding the parties’ respective information technology infrastructure in order to obtain the necessary concession.

### d. Categories or Types of ESI

A more difficult issue that should be discussed at the Rule 26(f) Conference is the categories of records to be preserved and produced. It may be possible to agree upon the categories of records to be preserved. Of course, this may be difficult to achieve at the initial Conference because the parties may not have sufficient information to reach an agreement on the potential categories of documents. The parties, however, are generally equipped to agree on the types of ESI that will be subject to discovery. This should include an agreement on specific file extensions that will or will not be within the scope of discovery.

### e. Search Terms

Many commentators and some courts have suggested that use of search terms is an appropriate mechanism to limit the burdens of e-discovery. With enterprise search and retrieval technology advances over the last few years, use of search terms is a realistic way to limit the scope of e-discovery. These tools, however, are not inexpensive and may not be readily available within your computing environment. Moreover, the ability to reach an agreement on topics or search terms is dependant upon the level of cooperation of the parties, and whether it is practical at the early stages in the case to identify the potentially relevant search parameters. Finally, search terms parameters have to be carefully crafted and multiple levels of searches may be required to identify truly relevant ESI. For example, a string search of 200 terms may recall so many records that even if the parties were able to agree on those terms, the results of the search are so massive that use of search terms in that context may not reduce or eliminate the burden.

## 2. Accessibility of Data

Federal Rule 26(b)(2)(B) now creates two tiers of ESI: (i) accessible data; and (ii) “not reasonably accessible” data. ESI that is not reasonably accessible need not be produced in the first instance, absent a showing of “good cause” by the requesting party. It is important to note that Rule 26(b)(2)(B) is a rule of production, **not preservation**. Accordingly, preservation obligations may apply to ESI, even if the ESI may not ultimately have to be produced based on inaccessibility. The Federal Rules do not define the specific types of data that are “not reasonably accessible.” Commentators generally agree that inaccessible data includes: (i) fragmented data; (ii) deleted data; and (iii) some back-up tapes.

While there is no obligation to either discuss or disclose at the Rule 26(f) Conference the categories of ESI that a party is claiming to be “not reasonably accessible” (such an obligation implicitly arises in the context of the Rule 34 response to request for the production of documents), we recommend that you use the Rule 26(f) Conference to attempt to reach an agreement, as to what data is “not reasonably accessible,” and the extent to which such data should be preserved. Counsel should be prepared to demonstrate why the data in question is “not reasonably accessible,” including discussing the costs and burdens associated with preserving, searching for and producing such data.

## 3. Production

Under Rule 26(f), the parties are required to discuss the scope of production and the production formal.



### **a. Scope of Production**

The scope of the document production should be limited through a discussion of the very same topics outlined above relating to preservation. It should be noted that the scope of preservation is likely to be broader than the scope of what is ultimately produced.

### **b. Production Format**

Production format is required to be addressed under Federal Rule 26(f). To understand this issue, it is necessary to explain the changes to Amended Rule 34(b) regarding procedures for requesting ESI and responding to requests for ESI.

Under Amended Rule 34(b), a request may “specify the form or forms in which electronically stored information is to be produced.” The responding party may object to the requested form. If the responding party objects to the requested form (or if the request does not specify the form), the responding party must state the form or forms that it intends to use. However, the responding party must produce the information in the form in which it is “ordinarily maintained” or in a “reasonably usable” form. In other words, like other forms of discovery, the production of ESI is subject to “requirements to protect against deliberate or inadvertent production in ways that raise unnecessary obstacles for the requesting party.” (Committee Notes to Amended Rule 34(b)).

Production format discussions raise a very hot issue – the production of data as TIFF/PDF images versus native productions. A native document is the document as it sits on the computer system with its associated metadata. In the past, responding parties have generally resisted producing documents in native format for several technical reasons: (i) production numbers cannot be affixed on each page; (ii) confidentiality designations cannot be affixed per page; and (iii) information cannot be redacted.

In hearings on the then proposed Amended Federal Rules, the Advisory Committee struggled with the format issue. It initially provided that ESI should be produced as “ordinarily maintained,” meaning native. There was a great deal of pushback during the hearings due to the technical limitations in producing native documents. For this reason, the Committee added the language in a “reasonably usable form” but it did not define what that means. The only guidance that the Committee provided is that “the option to produce in a reasonably usable form does not mean that a responding party is free to convert electronically stored information from the form in which it is ordinarily maintained to a different form that makes it more difficult or burdensome for the requesting party to use the information efficiently in litigation.” (Committee Notes to Amended Rule 34(b)). There was discussion during the hearings that it is not reasonable to take a record that is searchable in its native form and then produce a non-searchable TIFF/PDF image without any metadata. The Committee Notes reflect this discussion by noting, “If the responding party ordinarily maintains the information...in a way that makes it searchable by electronic means, the information should not be produced in a form that removes or significantly degrades this feature.” (Committee Notes to Amended Rule 34(b)).

Format of production is going to be a significant issue to be addressed and resolved at the Rule 26(f) Conference. In conducting these discussions, it is important to bear in mind that some federal courts will not permit (as they have in the past) a producing party to only produce a non-searchable TIFF/PDF version of ESI. In that case, at least some metadata, and possibly an extracted text searchable version of the document, may have to be produced. In addition, courts are likely to require that at least some ESI be produced in native form because of the inability to read or manipulate the data in a TIFF/PDF format, even one that is searchable.

## **4. Inadvertent Disclosure of Privileged Material**

The revised Rule 26(f) also requires parties to include in the proposed discovery plan, “any issues relating to claims of privilege or of protection as trial-preparation material, including – if the parties agree on a procedure to assert such claims after production – whether to ask the court to

include their agreement in an order.” Accordingly, the parties should discuss whether or not they would like to propose to the court an agreed upon procedure for addressing inadvertently produced privileged documents. Any such agreement reached by the parties should provide that: (i) upon notification or discovery of an inadvertently produced document, the opposing party should be notified; (ii) all copies of the inadvertently produced document should be returned to the producing party or be destroyed; (iii) any copies of the inadvertently produced privileged document provided to third parties, contained in the discovery record, or filed with the court should be retried and, again, returned to the producing party or destroyed; and (iv) to the extent the receiving party is challenging the assertion of privilege, the receiving party may retain one copy of the inadvertently produced record until the issue of privilege can be timely addressed by the court, and during such time, the inadvertently produced documents shall not be used for any purpose other than challenging the privilege assertion. The discussion should also include whether the parties shall agree that the inadvertent production of ESI shall not constitute a waiver of privilege and, if so, under what circumstances would constitute a waiver.

### **5. Database/Application Discovery**

Discovery of databases or data generated from larger applications is a complex issue that will require further discussion among the parties, and may have to be addressed separately from the discussion of other types of ESI such as e-mail or MS Office documents. In many instances, the database or ESI generated from the application is too large to simply produce the entire database/application. In many other instances, the database/application is proprietary and, thus, the opposition could not read the ESI without the proprietary database/application.

As a general matter, it may be helpful for the parties to exchange a list of the types of data they seek that may emanate from a database or proprietary application. In some instances, it may make sense to exchange a list of cells contained in the database/application or sample reports that can be generated from the database/application.

### **6. Timing of Production**

Depending upon the size of the Legal Matter and the amount of ESI involved, the collection processing and production of ESI can take considerable time. Timing of the production should be addressed at the Rule 26 Conference. Don't over promise. The process of collecting, processing, reviewing and producing ESI takes longer than you think.

### **7. Cost Shifting**

As a general matter, we believe that it is difficult for a large corporation to successfully shift the cost shifting relating to the discovery of ESI, with the exception of data that is “not reasonably accessible” or in the face of overbroad and unreasonable ESI discovery requests. There is, however, no harm in asking for cost shifting and certainly using cost shifting as a means to rein in overzealous e-discovery fishing expeditions.

### **8. Open Issues to Address at the Federal Rule 16(b) Discovery Planning Conference**

It is important to agree on the areas of disagreement. In this way, the issues to be addressed in a Rule 16(b) Discovery Planning Conference can be clearly identified and teed up for the court, assuming the court is willing to address them.

## VII. CONCLUSION

As our world becomes dominated by ESI that is stored on distributed computing systems, we need to be more thoughtful about how we litigate complex cases. That is what the changes to the Federal Rules and Rule 26(f) are all about: taking a more thoughtful and surgical approach to litigation so we can focus on what needs to be proved, how to prove it, and doing it in an efficient manner. Rule 26(f) is also about bringing together a diverse team of attorneys, record and information managers, and corporate compliance officers to prepare for and facilitate the meet and confer process, which can no longer be limited to the over-exacting outside litigator. Most importantly, the amended Rule 26(f) is about a fundamental change in dynamics – adversaries working together to facilitate the best interests of their clients by reducing the risks and costs of e-discovery. Those lawyers that adapt to this change will best serve the interests of their clients by applying some of the practices outlined above.