

2007

In Pursuit Of Frcp 1: Creative Approaches To Cutting And Shifting The Costs Of Discovery

Mia Mazza

Emmalena K. Quesada

Ashley L. Sternberg

Follow this and additional works at: <http://scholarship.richmond.edu/jolt>

 Part of the [Evidence Commons](#)

Recommended Citation

Mia Mazza, Emmalena K. Quesada & Ashley L. Sternberg, *In Pursuit Of Frcp 1: Creative Approaches To Cutting And Shifting The Costs Of Discovery*, 13 Rich. J.L. & Tech 11 (2007).

Available at: <http://scholarship.richmond.edu/jolt/vol13/iss3/4>

This Article is brought to you for free and open access by UR Scholarship Repository. It has been accepted for inclusion in Richmond Journal of Law and Technology by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

**IN PURSUIT OF FRCP 1: CREATIVE APPROACHES TO
CUTTING AND SHIFTING THE COSTS OF DISCOVERY OF
ELECTRONICALLY STORED INFORMATION**

*Mia Mazza, Emmalena K. Quesada, and Ashley L. Sternberg**

Cite as: Mia Mazza, Emmalena K. Quesada, & Ashley L. Stenberg, *In Pursuit of FRCP 1: Creative Approaches to Cutting and Shifting Costs of Discovery of Electronically Stored Information*, 13 RICH. J.L. & TECH. 11 (2007), <http://law.richmond.edu/jolt/v13i3/article11.pdf>.

The most important rule of all is the last sentence of [FRCP] 1, which provides that the Federal Rules of Civil Procedure ‘shall be construed to secure the just, speedy, and inexpensive determination of every action.’ It is this command that gives all the other rules life and meaning and timbre in the realist world of the trial court.¹

I. INTRODUCTION

[1] The primary purpose of the Federal Rules of Civil Procedure (“FRCP”) is “to economize time and expense in trying ... suits.”² That

* Mia Mazza is a partner, Emmalena K. Quesada is an associate, and Ashley L. Sternberg is a former associate, all in the San Francisco office of Morrison & Foerster, LLP. They gratefully acknowledge the contributions of Geoffrey Graber, Eric Brooks, Elizabeth J. Miles, Michael Sachdev, and Christopher Jeu to this article. Mia also acknowledges her law partner Steven M. Kaufmann for his early teachings on the importance of FRCP 1.

¹ *In re Paris Air Crash*, 69 F.R.D. 310, 318 (C.D. Cal. 1975) (quoting FED. R. CIV. P. 1).

² *Warnick v. Bethlehem-Fairfield Shipyard, Inc.*, 68 F. Supp. 857, 858 (D. Md. 1946). *See also* *Des Isles v. Evans*, 225 F.2d 235, 236 (5th Cir. 1955) (stating that the primary purpose of the rules is to secure “speedy and inexpensive justice in a uniform and well ordered manner.”); *Marin v. Knopf*, 1 F.R.D. 436, 437 (S.D.N.Y. 1940) (“The real purpose of the Federal Rules of Civil Procedure . . . is to eliminate technicalities, delays

mandate is articulated in FRCP 1, which requires that the FRCP “be construed and administered to secure the just, speedy, and inexpensive determination of every action.”³ In reality, few parties to litigation in federal court receive the prompt and economical resolution that FRCP 1 seems to promise. Nowhere is this more evident than in cases involving the discovery of large volumes of electronically stored information.⁴

[2] Discovery of electronic information is now an everyday fact of litigation in the U.S. Advances in computer software and hardware (e.g., e-mail, instant messaging, voicemail, blogs, laptops, .pdfs, PDAs, zip or flash drives, databases, and network servers) have greatly increased the ability to generate, replicate, circulate, and accumulate electronic information. At the same time, however, those technological advances also have resulted in the exponential growth of electronically stored information (“ESI”) that may be relevant to, and is now routinely requested in, litigation. It is estimated that more than 90% of all information today is created and retained in an electronic format.⁵

and expense and secure prompt and effective adjudication on the merits when a cause of action is set forth.”); *Boysell Co. v. Colonial Coverlet Co.* 29 F. Supp. 122, 124 (E. D. Tenn. 1939) (“[T]he intention of the new rules is to expedite litigation, to save cost and principally and primarily to reach justice by obtaining a full disclosure of the truth in connection with any controversy.”).

³ FED. R. CIV. P. 1. Throughout this article, references to “FRCP,” “Rules,” or “Rule” are to the Federal Rules of Civil Procedure. References to “Amendments,” “Amendment,” or “Amended” are to the FRCP amendments that were effective December 1, 2006.

⁴ The authors are of the opinion that the ubiquitous term “e-discovery” will eventually, if not soon, become outdated. The discovery of electronically stored information (“ESI”) is fast becoming the norm, with companies often converting paper records into electronic format for storage, at times complete with semi-searchable text created through Optical Character Recognition (“OCR”) or other technology. Accordingly, in time, the term “e-discovery” should be retired so that litigants can go back to using the term “discovery.”

⁵ THE SEDONA PRINCIPLES: BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION 4 (Jonathan M. Redgrave et al. eds., 2005), available at

<http://www.thesedonaconference.org/content/miscFiles/RetGuide200409.pdf> [hereinafter THE SEDONA PRINCIPLES]; see also, e.g. *Byers v. Ill. State Police*, No. 99 C 8105, 2002 WL 264004, at *10 (N.D. Ill. May 31, 2002) (“E-mails have replaced other forms of communication besides just paper-based communication. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via e-mail”).

[3] Unfortunately for litigants, “the advantages of computerization in the business world—the reduction of costs and greater efficiency in recording, storing, manipulating and retrieving information—have not translated into advantages in discovery.”⁶ The explosive growth of ESI has changed the very nature of discovery, with new electronic complexities making the preservation and production of evidence far more challenging.⁷ It is an accepted fact that “the discovery of computer-based information [can] cost more, take more time and create more headaches than conventional, paper based discovery.”⁸

[4] There are several characteristics of ESI that can make its discovery particularly expensive and burdensome:

- It is dynamic, and thus it can be altered or destroyed by the ordinary operation of a computer, often without the operator’s knowledge or direction.
- It is persistent, and thus it can continue to exist despite apparent “deletion”—yet in a state that makes it difficult to locate, retrieve, or search.

⁶ Kenneth J. Withers, *The Real Cost of Virtual Discovery*, 7 FEDERAL DISCOVERY NEWS, Feb. 12, 2001, at 3.

⁷ See *Kemper Mortgage, Inc. v. Russell*, No. 3:06-cv-042, 2006 WL 2319858, at *2 (S.D. Ohio April 18, 2006) (“Computers have become a standard tool of doing business, with many associated benefits and costs. One of the benefits but also burdens is that it is easier to preserve a great deal of information than it was with paper systems. One of the unexpected costs of using the electronic tool is that it may become costly to abide by one’s duty to preserve evidence.”); *Wiginton v. CB Richard Ellis*, 229 F.R.D. 568, 572 (N.D. Ill. 2004) (“As contrasted with traditional paper discovery, e-discovery has the potential to be vastly more expensive due to the sheer volume of electronic information that can be easily and inexpensively stored on backup media Depending on how the electronic data is stored, it can be difficult, and hence expensive, to retrieve the data and search it for relevant documents.”) (citations omitted).

⁸ Withers, *supra* note 6, at 3. See, e.g., *Medtronic Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550, 558 (W.D. Tenn. 2003) (“[T]he cost of restoring, de-duplicating, and designing and conducting a search of all 996 backup tapes reasonably could be in the range of several million [dollars] . . . [not including] the costs of privilege review and actual production.”).

- It is dependent upon the technology that created it, and thus it can be incomprehensible when separated from that system.⁹
- It is used and retained in a variety of different forms and in a multiplicity of locations (including servers, backup tapes, desktops, laptops, PDAs, home computers, and so forth), many of which are hard to reach.

[5] Finally, the volume of ESI is staggering in comparison with conventional paper discovery. It is not infrequent for a terabyte (i.e., 1,000 gigabytes, 1 million megabytes, or 500 billion typewritten pages) of ESI to be at issue in large civil litigation or in government investigations.¹⁰ The sheer magnitude and diversity of ESI that must be dealt with creates significant difficulties and costs for lawyers and litigants.

[6] For example, the cost of responding to a discovery request can be in the millions of dollars if several years' worth of archived e-mails and files must be located, preserved, restored, sorted through, and collected in a forensically sound manner.¹¹ While companies retain vast amounts of ESI, much of it is stored for disaster-recovery purposes.¹² Therefore, "[r]etrieving computer based records or data is not the equivalent of getting the file from a file cabinet or archives."¹³ The cost of restoring certain types of electronic information to a readable, searchable, and reviewable form can be astronomical, particularly in light of the limited amount of relevant information that the exercise may ultimately yield. Finally, there is the time and cost to cull, de-duplicate, review, and cleanse the data to exclude or redact non-relevant or confidential material before production. Assuming it takes a skilled attorney using available technology an average rate of one hour to review 100 documents, it would

⁹ COMMITTEE ON RULES OF PRACTICE AND PROCEDURE, REPORT OF THE JUDICIAL CONFERENCE COMMITTEE ON RULES OF PRACTICE AND PROCEDURE 23 (Sept. 2005), available at <http://www.uscourts.gov/rules/Reports/ST09-2005.pdf>

¹⁰ See MANUAL FOR COMPLEX LITIGATION (4TH) § 11.446 (2004).

¹¹ Linda Volonino, *Electronic Evidence and Computer Forensics*, 12 COMM. ASS'N FOR INFO. SYS. 463 (2003), available at <http://cais.isworld.org/articles/12-27/article.pdf>.

¹² See, e.g., *Lipco Elec. Corp. v. ASG Consulting Corp.*, No. 8775/01 2004 WL 1949062, at *8 (N.Y. Sup. Ct. 2004) ("electronic records [often] are not stored for the purposes of being able to retrieve an individual document" but rather "for emergency uploading into a computer system to permit recovery from catastrophic computer failure").

¹³ *Id.*

take him or her 5 years to review 1 million documents working 2,000 hours per year.

[7] At every stage of discovery, decisions must be made about where to devote money and resources. Deep in the trenches of discovery and facing the task of collecting, reviewing, and producing millions of electronic documents, it is easy for attorneys to lose sight of the forest for the terabytes. With the mounting volumes of ESI being created and stored, the practical reality is that discovery must become “not just about uncovering the truth, but also about how much of the truth the parties can afford to disinter.”¹⁴ Parties, lawyers, and judges have an opportunity today to begin taking creative new approaches to the discovery of ESI, with a renewed focus on the important (yet often neglected) mandate of FRCP 1.

[8] This article discusses the directive of FRCP 1 (Section II) and articulates several practical approaches to cutting (Section III) and shifting (Section IV) the costs of the discovery of electronically stored information. The authors posit that the approaches discussed herein are consistent with the command and spirit of FRCP 1, under which the “[c]ost of discovery is a pertinent and appropriate consideration.”¹⁵

II. FRCP 1: THE PROMISE OF A JUST, SPEEDY, AND INEXPENSIVE DETERMINATION OF EVERY ACTION

[9] FRCP 1 is “one of the least frequently cited, but most important”¹⁶ of all the Federal Rules of Civil Procedure.¹⁷ It commands that the Rules “be construed and administered to secure the just, speedy, and inexpensive

¹⁴ *Rowe Entm’t, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 423 (S.D.N.Y. 2002).

¹⁵ *Butterworth v. Indus. Chem., Inc.*, 145 F.R.D. 585, 589 (N.D. Fla. 1991).

¹⁶ *United States v. Texas*, 523 F. Supp. 703, 723 (E.D. Tex. 1981); *see also* *N. River Ins. Co. v. Greater N.Y. Mut. Ins. Co.*, 872 F. Supp. 1411, 1412 (E.D. Pa. 1995) (calling FRCP 1 “important but often neglected”).

¹⁷ The Federal Rules of Civil Procedure were originally adopted in 1938 by the Supreme Court. It was the result of the work of an Advisory Committee appointed by the Supreme Court in 1935 pursuant to the Rules Enabling Act of 1934, ch. 651, § 1, 48 Stat. 1064 (1934) (current version at 28 U.S.C. § 2071-74 (2006)).

determination of every action.”¹⁸ Those three aims are “a basic tenet of modern dispute resolution, an article of faith in Anglo-American civil procedure.”¹⁹ They are “complementary” attributes that must be considered in every case.²⁰

[10] The Supreme Court has observed that FRCP 1 reflects a “national policy . . . to minimize the costs of litigation.”²¹ The goal of reducing litigation costs is underscored by the 1993 amendment to FRCP 1, which clarifies that the Rules must be “*administered*,” as well as “construed,” to secure “the just, speedy, and inexpensive determination of every action.”²² The Advisory Committee stated that “[t]he purpose of this revision . . . is to recognize the affirmative duty of the court to exercise the authority conferred by these rules to ensure that civil litigation is resolved not only fairly, but also without undue cost or delay.”²³

[11] The 1993 amendments to the FRCP were part of an ongoing effort throughout the 1980s and early 1990s to reduce the costs of litigation. For example, at the time the Advisory Committee was working on the 1993 amendments, Congress also was passing the Civil Justice Reform Act of 1990 (CJRA).²⁴ The CJRA was “rooted in more than a decade of concern

¹⁸ FED. R. CIV. P. 1; *see also, e.g.*, *Hanover Ins. Co. v. Liberian Oceanway Corp.*, 398 F. Supp. 104, 110 (D.P.R. 1975) (stating that courts are obligated “to construe the procedures in such a manner as to secure the fast, speedy and inexpensive determination of every action.”); *Kenney v. Cal. Tanker Co.*, 381 F.2d 775, 777 (3rd Cir. 1967) (“[Courts] have the responsibility to secure the just and speedy and inexpensive determination of every action”); *Brown v. Ford Motor Co.*, 57 F. Supp. 825, 827 (E.D. Mich. 1944) (“It is the duty of the courts to construe the Rules of Civil Procedure so as to secure the just, speedy and inexpensive determination of every action.”).

¹⁹ Patrick Johnston, *Problems In Raising Prayers to Level of Rule: The Example of Federal Rule of Civil Procedure 1*, 75 B.U. L. REV. 1325, 1326-27 (1995).

²⁰ *Boe v. Lane & Co.*, 428 F. Supp. 1179, 1183 (E.D. La. 1977).

²¹ *Farmer v. Arabian Am. Oil Co.*, 379 U.S. 227, 234 (1964).

²² FED. R. CIV. P. 1 (emphasis added); FED. R. CIV. P. 1, comm. n (1993). For a thorough discussion of the legislative history behind the 1993 Amendment to Rule 1, see Johnston, *supra* note 19, at 1325.

²³ FED. R. CIV. P. 1, comm. n (1993).

²⁴ Judicial Improvements Act of 1990, Pub. L. No. 101-650, 101-105, 104 Stat. 5089-98 (1990).

that cases in federal courts take too long and cost litigants too much.”²⁵ Under the CJRA, federal courts are required to implement a “civil justice expense and delay reduction plan” to ensure a just, speedy, and inexpensive resolution of civil disputes.²⁶

[12] To achieve the outcome that FRCP 1 contemplates, the FRCP must be construed “liberally.”²⁷ Nowhere is this more true than when the federal courts apply Rules governing discovery. FRCP 1 has been cited in denying discovery that “would properly be characterized as a fishing expedition, causing needless expense and burden to all concerned.”²⁸ Courts also have held that the sharing of information obtained in discovery with litigants in comparable cases is consistent with FRCP 1, as it eliminates the time and expense of rediscovery.²⁹ The directive of FRCP 1 is thus clear: “[t]he provisions governing discovery under the Federal Rules of Civil Procedure are more than mere procedural guidelines to be consulted at the pleasure of a party to a federal suit. The language of these rules is carefully drafted and specific in its terms in order that they ‘secure the just, speedy, and inexpensive determination of every action.’”³⁰

[13] According to the Advisory Committee, the 1993 amendment to FRCP 1 proclaimed the “central theme and purpose” of several contemporaneous Rule amendments, including “reducing excessive delays and expense in civil litigation; curtailing and eliminating frivolous claims and defenses; reducing burdens on litigants; and preserving scarce judicial resources.”³¹

²⁵ James Kakalik, et al., *Just, Speedy and Inexpensive? An Evaluation of Judicial Case Management Under the Civil Justice Reform Act*, INST. FOR CIVIL JUST. 1 (1996).

²⁶ 28 U.S.C. § 471 (1994).

²⁷ *Plant Econ., Inc. v. Mirror Insulation Co.*, 308 F.2d 275, 278 (3d Cir. 1962).

²⁸ *N. River Ins. Co. v. Greater N.Y. Mut. Ins. Co.*, 872 F Supp 1411, 1412 (E.D. Pa. 1995).

²⁹ *See, e.g., Baker v. Liggett Group, Inc.*, 132 F.R.D. 123, 126 (D. Mass 1990) (holding that sharing discovery is consistent with FRCP 1 and particularly appropriate in tobacco tort cases in which individual plaintiffs must litigate against large, corporate defendants); *Wauchop v. Domino's Pizza, Inc.*, 138 F.R.D. 539, 546 (N.D. Ind. 1991) (explaining that the sharing of discovery materials furthers the goals of FRCP 1 by eliminating time and expense involved in discovery).

³⁰ *Philpot v. Philco-Ford Corp.*, 63 F.R.D. 672, 675 (E.D. Pa. 1974) (citation and internal quoted omitted).

³¹ *Johnston*, *supra* note 19, at 1328.

III. PUTTING TECHNOLOGY TO WORK FOR YOU: INNOVATIVE STRATEGIES
FOR CUTTING THE COSTS OF ESI DISCOVERY

[14] Even the 1993 Advisory Committee could not have imagined how prescient and relevant was its emphasis on cost-efficiency, or how far-reaching the scope of its mandate, in view of the explosion of electronic data that was about to happen. While “20 years ago PCs were a novelty and e-mail did not exist,” by the mid-1990’s most companies were fully embracing e-mail and other forms of ESI as essential components of business communications and operations.³² Today, e-mail is a fundamental and economical form of internal and external communication for many organizations.³³ By some estimates, “more than 90 percent of all information is created in an electronic format.”³⁴ Moreover, the growth rate of ESI is likely to accelerate rapidly over the next decade.³⁵

[15] As the veritable tidal wave of ESI hits, companies are learning of its dark side through the often painful experience of trying to harness it for use and production in litigation. In theory, at least, many of the problems caused by advances in technology also can be solved or at least minimized by those same advances.³⁶ Electronic data can be difficult to manage, but it “can also greatly reduce the costs of discovery and facilitate the pretrial preparation process.”³⁷ When properly employed, “electronic discovery allows a party to organize, identify, index, and even authenticate

³² THE SEDONA PRINCIPLES, *supra* note 5, at 1.

³³ *See, e.g.*, Byers v. Illinois State Police, No. 99 C 8105, 2002 WL 1264004, at *10 (N.D. Ill. May 31, 2002) (“E-mails have replaced other forms of communication besides just paper-based communication. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via e-mail.”)

³⁴ THE SEDONA PRINCIPLES, *supra* note 6, at 1.

³⁵ Richard K. Herrmann, Vincent J. Poppiti & David K. Sheppard, *Managing Discovery in the Digital Age: A Guide to Electronic Discovery in the District of Delaware*, 8 DEL. L. REV. 75, 75 (2005).

³⁶ David K. Isom, *Electronic Discovery Primer for Judges*, 2005 FED. CTS. L. REV. 1, at II. H. 1 (Feb. 2005).

³⁷ RICHARD VAN DUIZEND, WORKING GROUP ON ELECTRONIC DISCOVERY, CONFERENCE OF CHIEF JUSTICES, GUIDELINES FOR STATE TRIAL COURTS REGARDING DISCOVERY OF ELECTRONICALLY-STORED INFORMATION iv (Review Draft, Sept. 2005).

documents in a fraction of the time and at a fraction of the cost of paper discovery while virtually eliminating costs of copying and transport.”³⁸

[16] To that end, in recent years companies and counsel have taken advantage of one of the positive features of electronic data: its searchability. One of the most effective ways to reduce a large collection of data is through an electronic culling and vetting process.³⁹ ESI search technologies and software allow litigants to “search through far more documents than human beings could hope to review manually.”⁴⁰ While courts have, to some extent, endorsed search technology in discovery, the cases still lag far behind the available technology.

[17] This section suggests pragmatic, innovative approaches to reducing the costs of the ESI discovery process, in furtherance of the “just, speedy, and inexpensive” resolution of parties’ disputes. It focuses on the two most costly and burdensome aspects of ESI discovery: document preservation and attorney review. It also briefly discusses the need for companies to implement rigorous controls over data lifecycle management so that their electronically stored information does not become so voluminous as to be unmanageable when litigation hits. With these approaches, companies and counsel can make use of the special features of ESI to work toward fulfilling the mandate of FRCP 1.

[18] Even with the cost-cutting measures that search technologies provide, however, the fact remains that the more and more ESI there is to search, “the more expensive it is to discover all the relevant information.”⁴¹ To efficiently and effectively litigate in the electronic age, requesting parties should focus their search efforts on identifying and producing the most important portions of accessible and relevant ESI, rather than undertaking the Herculean task of trying to find and produce *all* responsive ESI. Conversely, requesting parties must give up the notion that they are entitled to every shred of relevant evidence, no matter how tangential

³⁸ *Id.* at iv-v.

³⁹ Robert Douglas Brownstone, *Collaborative Navigation of the Stormy E-Discovery Seas*, 10 RICH. J.L. & TECH 53 (2004).

⁴⁰ THE SEDONA PRINCIPLES, *supra* note 5, at 6.

⁴¹ *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 311 (S.D.N.Y. 2003) (*Zubulake I*).

A. KEYWORD SEARCHING

1. THE DEVELOPMENT OF KEYWORD SEARCHING IN DISCOVERY

[19] Long before the discovery of electronically stored information became part of everyday litigation, parties employed culling techniques during the document collection and pre-review processing stages of litigation to winnow down the overall volume of documents potentially relevant to a given litigation. Typically, those filters involved date parameters, custodians, and file types unlikely to contain relevant information. But given the exponentially greater (and ever-increasing) volume of ESI used and retained by companies in recent years, those traditional restrictions are no longer proving sufficient: the amount of data left to review is still often prohibitively colossal, sometimes numbering in the terabytes. The “cost and time required to have legal professionals read documents closely” for responsiveness, privilege, and other confidentiality concerns, especially “in the context of cases involving hundreds of thousands (or even millions) of pages of records, can be astronomical.”⁴² Moreover, with the continued proliferation of vast quantities of ESI, there is an ever-increasing likelihood that responsive documents will not be located and “that some privileged records may slip through the review process.”⁴³

[20] To minimize these problems, parties and their counsel can take advantage of one of the more positive features of electronic data: “because it can be searched automatically, key words can be run for privilege checks” and to identify potentially responsive documents.⁴⁴ “Keyword searching can, at least in theory, assist in all aspects of e-discovery.”⁴⁵

[21] As discussed herein, a keyword search is a method for culling down a large collection of ESI. The producing party searches the collection for certain words, dates, or other parameters specified by a user and then—in

⁴² Steven C. Bennett, *E-Discovery by Keyword Search*, 15 PRAT. LIT. 3, 9 (May 2004).

⁴³ *Id.*

⁴⁴ *Zubulake I*, 217 F.R.D. at 318.

⁴⁵ Bennett, *supra* note 42, at 10.

the most common application today—eliminates from attorney review those documents that do not contain those parameters. Keyword searching “permits a party to search” a collection of documents “more efficiently.”⁴⁶ Narrowing a pool of data by using keywords, such as names, e-mail addresses, and business terms relevant to the subject matter of the litigation, allows for a “targeted and focused discovery search.”⁴⁷

[22] Parties should consider how keyword searches also may be used beyond the context of the culling down of data before attorney review. A “significant benefit” of keyword searching is the potential “reduction in the burden and cost of electronic discovery (both in terms of locating and preserving relevant records, and in terms of identifying potentially privileged or confidential records).”⁴⁸ Assuming that a company has or is willing to acquire (through assistance of a vendor or otherwise) technology that allows it to search for keywords in ESI, protocols can be established to develop and deploy search terms to assist in the process of identifying ESI to be preserved and/or collected for a particular litigation. Keywords also could be used to facilitate a more efficient and effective search for privileged, private, and confidential information in ESI.

[23] Keyword searches are appropriate for identifying potentially privileged documents.⁴⁹ At a minimum, keyword searching can immediately identify documents that contain the words “privileged” or “confidential.”⁵⁰ A more sophisticated deployment of keyword searching could locate documents containing the names of relevant counsel, or other case-specific sensitive terms.⁵¹ Another appropriate situation to use keyword searching is where opposing counsel demands production of information on a specific subject.⁵² Keyword searching of ESI can also be particularly helpful where opposing counsel demands production of

⁴⁶ *Windy City Innovations, LLC v. American Online, Inc.*, No. 04 C 4240, 2006 WL 2224057, at *3 (N.D. Ill. July 31, 2006).

⁴⁷ *Reino de Espana v. Am. Bureau of Shipping*, No. 03 Civ. 3573 LTS/RLE, 2006 WL 3208579, at *3 (S.D.N.Y. Nov. 3, 2006).

⁴⁸ Bennett, *supra* note 42, at 12.

⁴⁹ *Id.*; *see also Zubulake I*, 217 F.R.D. at 318.

⁵⁰ Bennett, *supra* note 42, at 12.

⁵¹ *Id.*

⁵² *Id.* at 13.

information on a specific subject. Applying traditional document production techniques to that situation would require extensive, time-intensive, hands-on review of individual records one-by-one to identify responsive documents.⁵³ A keyword search, by contrast, would greatly reduce the time and costs necessary to locate those documents.⁵⁴

2. ENDORSEMENT OF KEYWORD SEARCHING AND FILTERING BY THE COURTS

[24] There is no real dispute that keyword searching, along with traditional filters such as date, custodian, and file type, can streamline the production of ESI.⁵⁵ Courts increasingly have endorsed the use of keyword searches and filters as a necessary component of discovery.⁵⁶ The use of search terms to reduce the volume of ESI at issue in a particular litigation matter “strikes a reasonable balance” between a requesting party’s needs and a producing party’s burden.⁵⁷ “Indeed, a principle advantage of electronic information is that high-speed methods exist to determine the existence of patterns of words, thereby allowing the narrowing of searches for relevant information.”⁵⁸

[25] A growing body of law acknowledges and endorses the use of “performing a key word search,” to cull down the volume of ESI to be reviewed by attorneys, to “control costs” in the review and production of

⁵³ *Id.*

⁵⁴ *Id.*

⁵⁵ *See, e.g., In re CV Therapeutics, Inc., Sec. Litig.*, No. C-03-3709 SI 2006 WL 2458720, at *2 (N.D. Cal. Aug. 22, 2006) (“Employment of search terms is a reasonable means of narrowing [production.]”); *FTC v. Ameridebt, Inc.*, No. C-058025JSWEDL, 2006 WL 618563, at *1 (N.D. Cal. Mar. 13, 2006) (noting that “e-mail could likely be screened efficiently through the use of electronic search terms that the parties agreed upon”).

⁵⁶ Exceptions to this rule have related to circumstances where keyword searching was not feasible due to the nature of the storage medium. Yet notably, even these cases recognize that if keyword searching had been feasible, it would have been appropriate. *See, e.g. Commonwealth v. Ellis*, No. 97-192, 1999 WL 815818, at *11 (Mass. Aug. 27, 1999).

⁵⁷ *In re CV Therapeutics*, 2006 WL 2458720, at *2.

⁵⁸ THE SEDONA PRINCIPLES, *supra* note 5, at 44. *See also Lombardo v. Broadway Stores, Inc.*, No. G026581, 2002 WL 86810, at *8 (Cal. App. Jan. 22, 2002) (“Broadway urges the hard copy payroll documents were the same as the computerized data. Not so. The hard copy may have contained the same information, but the information was not equally accessible.”).

documents.⁵⁹ The increasing familiarity of litigants with search technology—and with keyword searching in particular—has resulted in its adoption as an integral part of discovery in many cases. In *Medtronic Sofamor Danek, Inc. v. Michelson*, the court noted that “[p]roducing electronic data requires, at a minimum, . . . designing and applying a search program to identify potentially relevant electronic files.”⁶⁰ Commentators proffer that “[n]o sophisticated party or attorney seriously contends that an electronic vetting process is unnecessary.”⁶¹

[26] Case law makes it clear, however, that while keyword searches may be employed to more efficiently identify potentially relevant documents and cull down collections in a particular litigation, that should be done only after careful consideration of what keywords should be used.⁶² When keywords are properly developed, some courts are enthusiastic about the use of search terms in document review and production, deeming the procedure fair, efficient, and reasonable.⁶³ It has also become common for courts and parties alike to fashion unique, individualized discovery protocols incorporating the use of keyword terms.⁶⁴

⁵⁹ *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 576 (N.D. Ill. 2004).

⁶⁰ *Medtronic Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550, 552 (W.D. Tenn. 2003).

⁶¹ *Brownstone*, *supra* note 39, at 26.

⁶² *See, e.g., Balboa Threadworks, Inc. v. Stucky*, No. 05-1157-JTM-DWB, 2006 WL 763668, at *5 (D. Kan. Mar. 24, 2006) (noting the importance of formulating a detailed search protocol).

⁶³ *Tulip Computers Int’l B.V. v. Dell Computer Corp.*, No. CIV. A. 00-981-RRM, 2002 WL 818061 (D. Del. Apr. 30, 2002)

⁶⁴ *Id.*; *see also Rowe Entm’t, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 433 (S.D.N.Y. 2002). In *Rowe*, in light of the court’s decision to shift the costs of the discovery of e-mails to the plaintiffs, the court also decided to shift the responsibility for search, review, and even production of the e-mails to plaintiffs, the *requesting* party. The court directed plaintiffs’ counsel to formulate a search procedure for identifying responsive e-mails, including specific word searches. *Id.* Plaintiffs were also responsible for reviewing the documents elicited by the search, identifying the e-mails they considered relevant, and providing those documents to defendants’ counsel in hard copy form with Bates stamps. Defendants were permitted an after-the-fact opportunity to review the documents produced in order to designate them for privilege and confidentiality. The court dismissed with privilege waiver concerns by holding that the fact that “a document has been reviewed by [opposing] counsel or by the expert shall not constitute a waiver of any claim of privilege or confidentiality.” *Id.*

[27] There is a key question, however, that has not yet been addressed in depth by courts: may a responding party act “unilaterally” in creating and deploying search terms to winnow down a pool of data, or must there be a “bilateral” agreement between parties regarding search criteria before action is taken?

[28] A good-faith unilateral approach to the development of keywords for filtering and culling ESI should be defensible in light of the well-established principle that it is the province of the party producing material to “determine what is responsive to discovery demands.”⁶⁵ As one court noted, this may “range from reading every word of every document to conducting a series of targeted key word searches,” but either way, “the producing party unilaterally decides on the review protocol.”⁶⁶ The review protocol to be unilaterally developed by the producing party logically should include “defining the set of data” to be reviewed by selecting “reasonable search criteria, including search terms.”⁶⁷ Sedona Principle No. 6 acknowledges the soundness of this approach: “[r]esponding parties are best situated to evaluate the procedures, methodologies and technologies appropriate for preserving and producing their own electronic data and documents.”⁶⁸

[29] Some courts have endorsed a unilateral approach to deciding upon and employing search terms.⁶⁹ Others, however, have required that the requesting party have an opportunity to provide “input . . . regarding proposed search terms.”⁷⁰ Interestingly, even those courts still acknowledge that the producing party, “as custodian[] of the computer

⁶⁵ THE SEDONA PRINCIPLES, *supra* note 5, at 31.

⁶⁶ *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280, 290 (S.D.N.Y. 2003) (*Zubulake III*).

⁶⁷ THE SEDONA PRINCIPLES, *supra* note 5, at 31.

⁶⁸ *Id.* at I, §6.

⁶⁹ *See, e.g. Benson v. St. Joseph Reg'l Health Ctr.*, No. H-04-04323, 2006 U.S. Dist. LEXIS 28795, at *14 (S.D. Tex. May 17, 2006) (denying a motion to compel a full search of documents and for explanation of how the producing party searched archived material for documents and e-mails responsive to the requests, the court held that it was “unnecessary for [the producing party] to explain the details of their method of searching, when they have certified and represented to the Court that they have complied fully with [the requesting party’s] requests and made reasonable efforts to find and disclose all responsive documents and emails”).

⁷⁰ *In re Priceline.com Inc. Sec. Litg.*, 233 F.R.D. 88, 91 (D. Conn. 2005).

files,” “shall be responsible for . . . sifting through the data for responsive information” and may do so “in the most efficient manner possible.” Thus while it may be permissible for a producing party to develop a list of keywords unilaterally to structure its search for responsive ESI, it may be preferable for that party to make some sort of effort to involve the requesting party in bilateral discussions before the search terms are deployed.

[30] *Treppel v. Biovail Corp.* is an example of this approach.⁷¹ In *Treppel*, the court first relied on a line of cases to conclude that, compared to “hard copy” document review and production, keyword searching to filter and cull down a document collection is “more appropriate in cases involving electronic data, where the number of documents may be exponentially greater.”⁷² Biovail, the producing party, initially sought the plaintiff’s input in “defining the scope of any review of electronic records by stipulating which files would be searched and what search terms would be utilized.” Treppel declined, however, “apparently believing that ‘the use of search terms has no application to the standard discovery process of locating and producing accessible hard copy and electronic documents.’”⁷³ The court deemed that a “missed opportunity” for the plaintiff, but still chided Biovail for failing to deploy its search protocol, noting that “[a]bsent agreement” with the plaintiff, “Biovail should have proceeded unilaterally” and produced “all responsive documents located by its search.”⁷⁴ The court then ruled that Treppel could weigh in with “any specific concerns about the scope of the search.”⁷⁵

[31] Likewise, in *In re CV Therapeutics, Inc. Sec. Litig.*, the court ordered the defendant to disclose to the plaintiff the search terms it used to reduce the “universe of extant electronic records in this case,” to give the plaintiff

⁷¹ *Treppel v. Biovail Corp.*, 233 F.R.D. 363 (S.D.N.Y. 2006).

⁷² *Id.* at 374 (citing a line of cases approving keyword searches, including *United States ex rel. Tyson v. Amerigroup Ill., Inc.*, No. 02 C 6074, 2005 WL 3111972, at *2 (N.D. Ill. Oct. 21, 2005) (approving of parties’ stipulated search terms to review e-mails)).

⁷³ *Treppel*, 233 F.R.D. at 374.

⁷⁴ *Id.* at 374.

⁷⁵ *Id.* at 375.

“an avenue to test or assess the scope of the search terms” used.⁷⁶ When the plaintiff contested the adequacy of the search terms, however, the court held that the defendant’s unilateral development and use of keyword searches was a “reasonable means of narrowing the production in this instance” in light of the requesting party’s failure to “set forth an alternative search methodology” or otherwise bring any “specific challenge to the search terms.”⁷⁷

[32] Many cases are even more explicit about requiring parties to negotiate a mutually agreeable keyword list for filtering and culling as part of the discovery planning process, rather than proceed unilaterally.⁷⁸ The trend toward bilateral discussions regarding, and/or mutual agreement in, the use of keyword search methodologies may find additional support in the recent amendments to the FRCP concerning discovery of electronically stored information, effective December 1, 2006 (“Amendments”), which place emphasis on addressing discovery protocols openly, and as early as possible, in the litigation.⁷⁹

[33] There are, in sum, no bright-line rules regarding the development and use of keyword searches by a producing party. It may well be that as the volume of ESI exponentially increases and it becomes more and more infeasible to conduct attorney review of ESI without first substantially culling down the data, keyword searching will become a standard component of litigants’ joint discovery plans. In the meanwhile,

⁷⁶ *In re CV Therapeutics Sec. Litig.*, No. C-03-3709 SI 2006 WL 2458720, at *2 (N.D. Cal. Aug. 22, 2006).

⁷⁷ *Id.*

⁷⁸ *See, e.g.* *Balboa Threadworks, Inc. v. Stucky*, No. 05-1157-JTM-DWB, 2006 WL 763668, at *5 (D. Kan. Mar. 24, 2006) (requiring parties to meet and confer in an attempt to agree on a formulation of appropriate search protocol, whether they would use keyword searches and/or other search procedures, and advising parties to “lean heavily on their respective computer experts in designing such a protocol”); *Analog Devices, Inc. v. Michalski*, No. 01 CVS 10614, 2006 WL 3287382 (N.C. Super. Ct. Nov. 1, 2006) (directing producing party to conduct a word search using agreed-upon terms); *United States ex rel. Tyson v. Amerigroup Ill., Inc.*, No. 02 C 6074, 2005 WL 3111972, at *2 (N.D. Ill. Oct. 21, 2005) (endorsing negotiated keyword terms to be used in the review and production of e-mails).

⁷⁹ *Brownstone*, *supra* note 39 (advocating pre-production collaboration regarding specific selection criteria to be used in searches of the electronic data set).

producing parties should be free to take the lead in selecting and deploying keyword terms,⁸⁰ keeping in mind that they may best follow the spirit and directives of FRCP 1 by being up-front with the opposing party regarding proposed search terms (and receptive to its input). Where that is not possible, the producing party should, in an organized, methodical, and defensible fashion, unilaterally develop and use a keyword list to reduce the costs and burdens associated with attorney review of ESI.⁸¹

3. MOVING BEYOND CASE LAW: USE OF KEYWORD SEARCHING IN DOCUMENT PRESERVATION

[34] While the use of keyword searching in discovery is generally accepted by courts, that judicial endorsement has, to date, mainly been limited to the pre-review culling phase. Long before any ESI is collected or culled, however, keyword searches may offer similar efficiency benefits for the preservation of potentially relevant ESI, which is fast becoming a very complex and costly stage of discovery.⁸²

(A) THE NEW LEGAL LANDSCAPE OF DOCUMENT PRESERVATION

[35] While the duty to preserve evidence in the face of pending or anticipated litigation is not new,⁸³ the explosion of ESI has deepened the impact of that duty. The proliferation of e-mail and other forms of ESI in

⁸⁰ THE SEDONA PRINCIPLES, *supra* note 5, at 31 (“There is no principled reason to require more intrusive efforts merely because the party seeking discovery is suspicious of the efforts undertaken by the producing party.”); *see also* McCurdy Group, LLC v. Am. Biomedical Group, Inc., 9 Fed. Appx. 822, 831 (10th Cir. 2001) (affirming denial of motion to compel production of hard drives based on that party’s expression of skepticism that all relevant and non-privileged documents had been produced).

⁸¹ *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 374 (S.D.N.Y. 2006).

⁸² THE SEDONA PRINCIPLES, *supra* note 5, at 6; *see also* Isom, *supra* note 36.

⁸³ A party on notice that documents and information in its possession are relevant to pending or anticipated litigation matter has a duty to preserve those documents and data. *See, e.g.*, *Wm. T. Thompson Co. v. Gen. Nutrition Corp.*, 593 F. Supp. 1443, 1455 (C.D. Cal. 1984) (“While a litigant is under no duty to keep or retain every document in its possession once a complaint is filed, it is under a duty to preserve what it knows, or reasonably should know, is relevant in the action, is reasonably calculated to lead to the discovery of admissible evidence, is reasonably likely to be requested during discovery, and/or is the subject of a pending discovery request.”); JAMIE S. GORELICK ET AL., DESTRUCTION OF EVIDENCE § 3.11, 93 n.125 (1989 & Cumulative Supp. 2006).

corporations has increased exponentially the volume of data potentially relevant to any litigation.⁸⁴ This presents a particular challenge for large corporations that are constantly litigating multiple actions across different areas of law and various company departments. Large companies have worked hard to keep up with their duty to preserve evidence, but it has been difficult and costly. Each time a new litigation matter arises, or an existing matter expands in scope, the company is expected to initiate or expand a “litigation hold” on any manual or automatic processes that might delete relevant information, while the disposal of all non-relevant information continues.⁸⁵ The decision of what to include in the scope of a “litigation hold” is thus of great importance.⁸⁶ Once the scope of a litigation hold has been determined, it is up to a party and its counsel to take reasonable steps to see that sources of information within the scope are located and actually placed on hold during implementation.⁸⁷

[36] There is no entirely risk-free approach to preserving ESI for litigation. While many companies traditionally have attempted to take a mechanical “save everything” approach to reduce the risk of spoliation, they now are realizing that saving “everything” is simply no longer feasible in light of the colossal volume of ESI, nor is it even a good idea. Over time, that approach eventually would compromise the company’s ability to meet its obligation to preserve and produce relevant evidence in litigation: as the volume of data being “held” for litigation grows to an

⁸⁴ Mia Mazza & Ashley Sternberg, “*Good Faith*” *Safe Harbor of FRCP 37(f) to Prevent Drowning in Zubulake?*, A.B.A. SEC. LITIG., COMM. ON TRIAL EVIDENCE 1 (Winter 2007).

⁸⁵ See, e.g. *Zubulake v. UBS Warburg LLC*, 220 F.R.D. 212, 218 (S.D.N.Y. 2003) (*Zubulake IV*) (“Once a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy and put in place a ‘litigation hold’ to ensure the preservation of relevant documents.”); *DaimlerChrysler Motors v. Bill Davis Racing, Inc.*, No. 03-72265, 2005 WL 3502172, at *1 (E.D. Mich. Dec. 22, 2005) (“[N]ormal procedures for destruction of documents must . . . be suspended when a party is on notice that they may be relevant to litigation, and the failure to make an adequate search of such documents before their destruction may be evidence of bad faith.”); *Rambus, Inc. v. Infineon Tech.*, 220 F.R.D. 264, 281 (E.D. Va. 2004).

⁸⁶ Mazza & Sternberg, *supra* note 84.

⁸⁷ See *Zubulake v. UBS Warburg LLC*, 229 F.R.D. 422, 432 (S.D.N.Y. 2004) (*Zubulake V*); see also *Phoenix Four, Inc. v. Strategic. Res. Corp.*, 446 F.Supp.2d 205 (S.D.N.Y. 2006).

unmanageable size, it becomes increasingly difficult to find the information that is “relevant” to any particular case.⁸⁸

[37] The need for a more efficient and effective approach to preservation has been driven home for companies by the increased rate at which requesting parties are seeking sanctions for alleged spoliation of ESI.⁸⁹ A range of negative consequences can follow from a company’s failure to preserve evidence, including monetary sanctions, evidence or issue preclusion, adverse inference jury instructions, and even outright dismissal.⁹⁰ One notorious example is *Zubulake v. UBS Warburg LLC*.⁹¹ In April 2005, a jury awarded damages of more than \$29 million against UBS Warburg (“UBS”) after the court ruled that UBS’s preservation efforts had failed to prevent the disposal of certain relevant, responsive documents and gave the jury an adverse inference instruction.⁹²

[38] In light of cases like *Zubulake* and the evolving complexity of ESI, companies are now taking more substantial steps than in the past to ensure proper document preservation and to avoid discovery sanctions. This means that companies and counsel must make critical decisions early on—often before any litigation is filed—about where potentially relevant ESI may be located, and what must be done to preserve it for litigation. For large companies that create high volumes of ESI and that are constantly in numerous litigation matters at once, compliance with (and concern about)

⁸⁸ See, e.g., *Zubulake IV*, 220 F.R.D. at 217 (holding that requiring a company to retain every shred of electronic and paper documents upon notice of potential litigation would “cripple large corporations” who are “almost always involved in litigation”).

⁸⁹ See *West v. Goodyear Tire & Rubber Co.*, 167 F.3d 776, 779 (2d Cir. 1999) (“Spoliation is the destruction or significant alteration of evidence, or the failure to preserve property for another’s use as evidence in pending or reasonably foreseeable litigation.”).

⁹⁰ See, e.g., *Silvestri v. Gen. Motors Corp.*, 271 F.3d 583, 593 (4th Cir. 2001) (holding a dismissal the appropriate sanction for plaintiff’s failure to discharged duty to prevent spoliation of evidence); *Trigon Ins. Co. v. United States*, 204 F.R.D. 277, 291 (E.D. Va. 2001) (holding that adverse inferences respecting substantive testimony and credibility of experts was the appropriate sanction for spoliation). Further, any knowing destruction of documents sought in litigation may be punished by lengthy prison sentences. See, e.g. 18 U.S.C. § 1512 (c).

⁹¹ *Zubulake V*, 229 F.R.D. at 422.

⁹² *Id.* at 439-40. See *Mazza & Sternberg*, *supra* note 84 (discussing *Zubulake*, new Rule 37(f), and their implications for future cases involving sanctions for spoliation).

the rigor now required in document preservation can be overwhelming. That has translated into a phenomenon of companies spending a great deal of time and money on the purchase and implementation of additional software, hardware, and other technology infrastructure, and the hiring of personnel tasked solely with the responsibility of preserving ESI for litigation, in an attempt to increase the likelihood that sources of potentially relevant ESI will be identified and retained for each new litigation matter that arises.⁹³ This costly practice simply cannot be reconciled with the mandates of FRCP 1.

[39] Rule 37(f), adopted with the FRCP Amendments, provides that sanctions under Rule 37 cannot be imposed 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,” absent exceptional circumstances.⁹⁴ The Advisory Committee note to Rule 37(f) does not shed any light on what qualifies as “the routine, good-faith operation of an electronic information system,” but it may signal a shift away from essentially expecting parties to do a perfect job of preserving ESI for litigation and toward a focus on whether parties have exercised “good faith” in their efforts to preserve potentially relevant ESI.⁹⁵

(B) KEYWORD SEARCHING AS PART OF A REASONABLE, GOOD-FAITH
APPROACH TO DOCUMENT PRESERVATION

[40] In light of the excessive costs and burdens now associated with document preservation, courts have acknowledged and held that companies need not preserve “every scrap of paper” in their files to comply with preservation obligations.⁹⁶ Companies can only strive to

⁹³ See THE SEDONA PRINCIPLES, *supra* note 5, at 14, cmt. 2.b (“[T]he non monetary costs,” including “the burdens on information technology personnel and the resources required to review documents” should also be considered). For example, large companies are now commonly finding it necessary to purchase a server whose sole purpose is to store preserved data for a particular litigation, which alone can cost nearly half a million dollars, or more. Telephone Interview with Discovery Vendor (Dec. 12, 2006).

⁹⁴ FED R. CIV PRO. 37(f)

⁹⁵ See Mazza & Sternberg, *supra* note 84.

⁹⁶ Danis v. USN Comm., Inc., No. 98 C 7482, 2000 WL 1694325, at *32 (N.D. Ill. Oct. 23, 2000).

make reasonable, good-faith efforts to preserve ESI for litigation, and that includes the process by which they locate potentially responsive ESI to be preserved.⁹⁷ *Zubulake* anticipated that companies would need to “be more creative” in the way they locate documents within the scope of their duty to preserve.⁹⁸ “In recognition of the fact that there are many ways to manage electronic data,” the court stated that “litigants are free to choose how [the task of preserving relevant documents] is accomplished.”⁹⁹

[41] Sedona Principle 11 also reflects the need for innovative solutions in the preservation context:

A responding party may satisfy its good faith obligation to preserve and produce potentially responsive electronic data and documents by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data most likely to contain responsive information.¹⁰⁰

The key factor is that companies undertake the preservation obligation in a reasonable, good-faith, well thought-out manner, in consideration of the facts of the litigation at issue.¹⁰¹ Decisions “relating to legal standards like ‘potentially relevant data’ are made every day by lawyers in a wide variety of contexts.”¹⁰² This includes companies and their counsel making judgment calls regarding which methods are reasonable and which are far

⁹⁷ Bennett, *supra* note 42, at 10.

⁹⁸ *Zubulake V*, 229 F.R.D. 422, 432 (S.D.N.Y. 2004).

⁹⁹ *Zubulake IV*, 220 F.R.D. 212, 218 (S.D.N.Y. 2003).

¹⁰⁰ THE SEDONA PRINCIPLES, *supra* note 5 at 39, prin. 11.

¹⁰¹ *Id.* at 27, cmt 6.b. (“[O]rganizations should define the scope of the data needed to appropriately and fairly address the issues in the case and to avoid unreasonable overbreadth, burden, and cost. Important steps in achieving the goal of reasonably limiting discovery may include collecting data from repositories used by key players rather than generally searching through the entire corporate computer system; defining the set of data to be collected by applying reasonable selection criteria, including search terms.”).

¹⁰² John L. Carroll & Kenneth J. Withers, *Observations on “The Sedona Principles,”* at 4 (Apr. 2003), available at <http://www.kenwithers.com/articles/sedona/observations.pdf>.

too burdensome to endure. When electronic records are voluminous, a keyword search should be employed where feasible.¹⁰³

[42] Through the use of keyword search technologies, a company may gain some assurance that it has made adequate good-faith efforts to comply with its obligation to locate and preserve potentially relevant information, while at the same time avoiding “the overwhelming burden of either preserving everything that exists in electronic form, or conducting a laborious page-by-page review of electronic records to determine what must be preserved.”¹⁰⁴

[43] *Zubulake* is in accord.¹⁰⁵ There, the court explicitly suggested that in fulfilling the preservation obligation counsel should consider using a system-wide keyword search to locate potentially relevant documents.¹⁰⁶ A copy of each “hit” could then be preserved and segregated to provide a pool of potentially relevant documents for later responses to discovery requests.¹⁰⁷

[44] Where keyword search technology is employed to facilitate document preservation, parties should consider adopting a more conservative approach than they would in the review stage of discovery. This is because the filtering out of documents that do not meet various keyword search criteria will be permanent. The parties will not have the ability to go back and redefine the scope of their search terms, as they would during the review stage, so companies and their counsel should choose objectively broad criteria to guard against potential spoliation claims. In other words, when using keywords for preservation purposes, search terms should be constructed with the intent of reaching ESI that is *potentially* relevant, and not just ESI that is relevant, in order to maximize defensibility.¹⁰⁸ (Later on, a narrower set of keyword search terms should be applied to cull down the data set to a volume that feasibly may be

¹⁰³ Bennett, *supra* note 42, at 10.

¹⁰⁴ *Id.* at 11.

¹⁰⁵ *Zubulake V*, 299 F.R.D. 422, 432 (S.D.N.Y. 2004).

¹⁰⁶ *Id.*

¹⁰⁷ *Id.*

¹⁰⁸ *See, e.g., Zubulake IV*, 220 F.R.D. 212, 217 (S.D.N.Y. 2003); *see also* Wm. T. Thompson Co. v. Gen. Nutrition Corp., 593 F. Supp. 1443, 1455 (C.D. Cal. 1984).

reviewed by attorneys.) To that end, companies should consider keyword searching as a tool to be used *in conjunction with* traditional preservation efforts such as identifying relevant custodians and locating the ESI that they created and reviewed during the relevant time period.

[45] In further support of defensibility, a party should make sure there is quality control and oversight in executing keyword searches. If it appears that a particular search is not returning responsive documents, a company should modify its search parameters or take other measures to ensure appropriate preservation. Also, a party should carefully document the search protocol employed. These, along with any other visible efforts a company can make to show that it made reasonable efforts to fulfill its preservation obligation in good faith, should go a long way toward warding off any future claims of spoliation, and position a party to invoke the protections in new FRCP 37(f).¹⁰⁹

4. LIMITATIONS OF KEYWORD SEARCHING

[46] While keyword searching has been embraced by courts, counsel, and parties alike, it is by no means an “e-discovery panacea.”¹¹⁰ There are challenging problems associated with using keyword searching, many of which are inherent to the available technology. Perhaps foremost among the limitations of keyword searching is the risk that the chosen search terms will result in a low recall rate of potentially relevant documents.¹¹¹

¹⁰⁹ See FED. R. CIV. P. 37(f) (“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.”).

¹¹⁰ Bennett, *supra* note 42, at 13.

¹¹¹ The recall rate “is a measure of the ability of a given retrieval methodology to find all of the potential responsive documents in a given collection.” Jason R. Baron, *Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used in E-Discovery*, 6 SEDONA CONFERENCE J. 237, 242 (2005). Recall and precision are important performance measurement terms in information retrieval. See Matt Deniston, *Concept Searching Whitepaper*, at 4 (Jan. 2003), http://www.litigationready.com/pdfFiles/concept_searching.pdf. (“Precision is a measure of the system’s ability to return only relevant documents from amongst all the documents in a given collection. It answers the question, ‘Is what I found relevant to what I was looking for?’ For example, if a search engine lists 80 documents found to match a query

Lawyers and experts struggle with formulating keyword search terms that will locate the largest possible number of potentially relevant documents without generating so many “hits” that the company essentially must preserve or review everything.

[47] First, people often do not use the same terms to describe the same idea. It can be very difficult, if not impossible in some circumstances, to identify every single word that relates to a relevant topic. Consequently, keyword searches often fail to be as broad as a particular concept requires. For example, a keyword search could include “car,” but neglect “automobile,” “Jeep,” or “wheels.” Because a keyword search is only as broad as the exact terms used, it is safe to assume that it will miss, or fail to recall, many relevant documents.

[48] Second, anyone who has ever used a Boolean search engine knows the importance and the difficulty of framing queries to maximize the recall rate of responsive documents. Building proper queries is a crucial step in the discovery process when keyword searches are employed to locate relevant documents. It “requires mastering the subtleties of ANDing and ORing terms in order to arrive at a list of documents worth reviewing.”¹¹² In addition, varying search engines often require slightly differing syntaxes for the construction of Boolean search terms. Accordingly, converting document requests into “a syntactically correct Boolean query” can challenge “even the most seasoned expert searcher.”¹¹³ In the words of one court, “[k]eyword searches are limited because they are literal and search only for an exact sequence of characters. Thus, they do not pick up

but only 20 of them are relevant, then the Precision would be 25%. Recall is a measure of how well the system can find all of the relevant documents in the database. It answers the question, ‘How much is out there?’ For example, there may be 100 matching documents, but a search engine may only find 80 of them. It would then list these 80 and have a recall of 80%.”) *Id.* at 3.

Id. The relationship between recall and precision “is an inverse one: increasing the recall rate invariably leads to a corresponding loss of precision, as more and more documents are retrieved to find the elusive remaining needle in the rest of the haystack.” *Baron*, at 4.

¹¹² Deniston, *supra* note 111, at 3.

¹¹³ *Id.*

variations or misspellings of words or names.”¹¹⁴ These kinds of errors easily can lead to protracted discovery disputes later on down the road.¹¹⁵

[49] Third, keyword searching naturally results in “false positives,” making any given search overinclusive from the outset. For example, the word “water” may also bring up search results that are totally unrelated to the meaning of the word you are searching for, such as “Water St.,” “watered down,” or “John Waters.” When dealing with a large universe of documents, it can be problematic when a keyword search generates “large numbers of nonresponsive records (along with responsive ones), all of which [need] to be sorted out through a labor-intensive manual review process.”¹¹⁶ This kind of result may, to some extent, defeat the efficiency rationales behind using a keyword search in the first place.¹¹⁷

[50] Finally, it is important for a company to recognize that in all likelihood it will need to perform multiple keyword searches across its various data systems, as equipment and platforms are not always compatible. “Unfortunately, there is no simple ‘universal translator’ for all this information.”¹¹⁸ When electronic information is maintained in different formats and stored in multiple locations, separate searches are typically required for each format and location.¹¹⁹ In many locations, keyword searching may not be feasible at all.¹²⁰

¹¹⁴ United States v. Triumph Capital Group, Inc., 211 F.R.D. 31, 50 (D. Conn. 2002).

¹¹⁵ See, e.g., Coleman Holdings, Inc. v. Morgan Stanley & Co., Inc., No. 502003CA005045XXOCAI, 2005 WL 679071, at *7 (entering default judgment against producing party for a host of discovery abuses, including failure to identify and timely disclose script errors in search terms that prevented it from locating responsive documents).

¹¹⁶ Baron, *supra* note 111, at 237 (describing in detail the problems arising from overbroad keyword searches in two cases).

¹¹⁷ *Id.*

¹¹⁸ Bennett, *supra* note 42, at 14.

¹¹⁹ *Id.*

¹²⁰ See, e.g., United States v. Triumph Capital Group, Inc., 211 F.R.D. 31, 50 (D. Conn. 2002) (“Keyword searches are also limited because they cannot be conducted on all files, such as image files that contain scanned documents or faxes.”). See also Public Citizen, Inc. v. Dep’t of Educ., 292 F. Supp. 2d 1, 7 (D.D.C. 2003) (“[B]ecause the database that was searched could not be relied upon to contain the requested information, it follows that no amount of searches of that database could be relied upon to turn up that information.”).

[51] Companies will need to account for the deficiencies associated with keyword searching and adapt their procedures accordingly. Deciding what to search for and executing the search must be the result of intensive strategy considerations. It is a challenging task, but is not insurmountable.¹²¹ The search terms employed “must be reasonably calculated to return relevant data.”¹²² It may even be necessary to engage an expert who could ensure that the search terms are properly constructed and are compatible with all relevant systems. If the search is not done properly, courts may order additional searches, which will increase the cost and burden of discovery.”¹²³

B. CONCEPT SEARCHING

[52] There can be no doubt that electronic search methodologies, as a means of culling down a data set for discovery, are here to stay. While keyword searching certainly has its place in winnowing down ESI in a cost-effective and speedy manner, and the technology currently is available through many vendors and software providers, it also leaves something to be desired in terms of both recall and precision.¹²⁴ Another, more recent development in search technology is concept searching, a promising alternative or complement to keyword searching.

[53] As of this writing, concept search tools have not been developed or marketed widely for integration into corporate computer systems and thus they typically are not used in-house for preservation or other initial stages of discovery. Concept search technology software, however, currently is available (at a price) for use in litigation to cull down and/or organize collections of ESI before attorney review.

¹²¹ Bennett, *supra* note 42, at 13.

¹²² THE SEDONA PRINCIPLES, *supra* note 5, at 54.

¹²³ *Id.*; see also *In re Amsted Indus., Inc.*, No. 01 C 2963, 2002 WL 31844956, at *2 (N.D. Ill. Dec. 18, 2002) (ordering additional, comprehensive searches after finding word searches already employed to be insufficient.).

¹²⁴ See Deniston, *supra*, note 111, at 14.

1. CONCEPT SEARCH TECHNOLOGY

[54] Concept search technology takes keywords to a whole new level. Unlike keyword search systems that match precisely what you type in and nothing more, “concept search systems strive to determine what you mean.”¹²⁵ Concept searching aims to improve upon the performance limitations of keyword searching by using sophisticated statistical and linguistic models to understand the meaning behind search terms. This is done by identifying word patterns and occurrences in documents, which are then translated into “concepts.”¹²⁶ The concept search tool then compares those “concepts” across the document set, looking for relationships between documents.¹²⁷

[55] Unlike traditional Boolean operators like “AND,” “OR,” and “!,” concept searching does not require any particular syntax or formatting. Users can enter words, sentences, even whole documents (such as document requests or complaints), and the technology will return a list of potentially related documents, often ranked by relevancy, or of custodians with the most “hits.”¹²⁸ By allowing lawyers to search for ideas as well as words, potentially relevant documents can be located even if specifically identified keywords do not appear in them. Conversely, concept search engines are “smart enough” to exclude irrelevant documents even if a potentially relevant “concept” does appear in them.¹²⁹ For example, if a reviewer entered “‘fired from job’ into a concept search engine, the engine is smart enough to exclude information such as flames, smoke and fireplaces. The concept search will also expand the search to dismissal, separation, layoff, suspension, etc.”¹³⁰ Essentially, concept searching is akin to “using a massive dictionary of the English language to find synonyms and related words.”¹³¹ Many concept search tools also operate much like a massive dictionary of the language used by a company and in

¹²⁵ *Id.* at 2.

¹²⁶ Kristin M. Nimsger & Michele C. S. Lange, *E is for Evidence: Examining Recent E-Discovery Developments*, 23 No. 2 GPSOLO 40, 44 (Mar. 2006).

¹²⁷ *Id.*

¹²⁸ Telephone Interview with Discovery Technology Vendor (Dec. 14, 2006).

¹²⁹ Deniston, *supra* note 109, at 2.

¹³⁰ *Id.*

¹³¹ Baron, *supra* note 111, at 240.

the larger document collection being searched, for example to identify code names for projects or nicknames for people at the company.

[56] There are several different theoretical approaches to finding data through concept searching. There are mathematical or statistical approaches,¹³² which often use probability applications to locate and rank related topics. Other approaches include linguistics or taxonomy.¹³³ Each has its own set of strengths and weaknesses.¹³⁴ Accordingly, most concept search software blends together some combination of the approaches to fit the needs of a particular litigation and to optimize precision and recall of relevant documents.¹³⁵ The goal is for the technology to come as close as possible “to modeling how humans parse text for semantic meaning.”¹³⁶

[57] When combined with traditional Boolean keyword searching, these search technologies provide a powerful tool for lawyers to substantially reduce the costs of discovery. The technologies allow litigants to decrease significantly the amount of material that must be reviewed by attorneys (or preserved in the first instance), and allow lawyers to identify and review the most relevant documents more quickly and accurately. “Theoretically, as technology improves, retrieving and searching data will become more standard and less costly.”¹³⁷

¹³² Deniston, *supra* note 111, at 4-9 (discussing different mathematical and statistical models used in concept search technology and stating that mathematical models are better at handling large data collections (measured in terabytes) than other approaches).

¹³³ Telephone Interview with contributing author to Sedona Conference’s Working Group on Methods for Search and Retrieval (Dec. 8, 2006). A taxonomy approach relies on an “if this, then that” approach. For example, with the word “pitch” you would retrieve general categories such as baseball, roofing, sales, etc. If you next selected “baseball,” you would get “fastball,” “curve,” “throw,” and so on, all clustered around baseball. A linguistic approach, in addition to focusing on context and circumstances of use of words in language, relies on statistical applications and also looks at the practices and activities in which the words and phrases are used.

¹³⁴ *Id.*

¹³⁵ Telephone Interview, *supra* note 126. Indeed, the concept search tools that many vendors use are a blend of mathematical, statistical, and linguistic approaches.

¹³⁶ Deniston, *supra* note 111, at 8.

¹³⁷ *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 572 (N.D. Ill. 2004).

2. HOW CONCEPT SEARCHING CAN AMPLIFY THE EFFICIENCIES OF KEYWORD SEARCHING

[58] Concept searching is able to do what keyword searching does, and more. Among experts, it is thought to be a better and more useful application in that it may offer parties a better chance than keyword searching of locating relevant documents.¹³⁸ A concept search will return documents that relate to the same idea as the query word, sentence, or paragraph, making it arguably a more reliable method of locating responsive documents.¹³⁹

[59] Keyword searching is a great “brute force” way to find exact matches.¹⁴⁰ If litigants need to identify documents containing certain names or terms, then keyword searching can be an essential and powerful tool to find that information.¹⁴¹ If, on the other hand, a party needs to identify responsive documents and has a less refined sense of which documents could be responsive, then concept searching might be a better option. For example, upon receiving a document request, rather than constructing and executing complex keyword searches, concept searching would let a party copy and paste the relevant part(s) of the request into the search software. Shortly thereafter the party would have a list of every conceptually-related document, including documents that didn’t contain any of the specific words that were “pasted” into the software.¹⁴²

[60] The use of concept searching in discovery is, to the knowledge of the authors, currently limited to attorney review and (in fewer instances) pre-review culling. As the technology becomes more widely available, however, it also could be useful in identifying documents to preserve for anticipated litigation. As discussed above, the preservation obligation presents a vexing challenge for large companies that, at the time the duty

¹³⁸ Telephone Interview with Discovery Vendor (Nov. 27, 2006); Telephone Interview with contributing author to Sedona Conference’s Working Group on Methods for Search and Retrieval (Dec. 8, 2006).

¹³⁹ *Id.* Most search technologies allow for further control of the returned results, such as de-duplication or grouping of emails by custodian or chain.

¹⁴⁰ Deniston, *supra* note 111, at 3.

¹⁴¹ *Id.*

¹⁴² *Id.* at 2.

to preserve arises, may not have received a complaint or preservation letter. Yet, they are nevertheless expected to identify and safeguard all information that may turn out to be potentially relevant.¹⁴³ Concept searching could be used as a tool or safeguard mechanism in this context by counsel, along with traditional preservation methods, to ensure the widest recall of potentially relevant documents. For example, counsel could paste phrases and paragraphs from a complaint, preservation letter, or newspaper article into the search tool, and the concept search would identify relevant documents and further detail which custodians had the greatest number of relevant document “hits.” This could serve as a starting point for preservation, or as a final sweep, depending on the situation.

[61] Using concept search technology instead of keyword searching is particularly appropriate in cases that cannot be easily reduced to specific terminology, or that are complex in nature. For example, in cases that center on wrongful behavior, such as sexual harassment, it can be challenging to identify search terms that adequately describe harassing behavior.¹⁴⁴ Concept searching could locate relevant ESI that relates to, rather than contains, specific words.

[62] Concept search technology could also be used to further minimize the risk that the most sensitive documents—those that are privileged, private, confidential, or otherwise protected from discovery—do not get into the hands of the opposing party. Some commentators suggest that the nature and volume of discovery in the era of electronic information make it likely that old attorney-client privilege issues will need to be examined in a new light.¹⁴⁵ As the volume of documents requested and produced increases, the results are twofold in the privilege context. First, privilege review is more costly and time-consuming than ever. Second, even with a comprehensive privilege review, the sheer volume of documents requested and produced increases the risk of inadvertent disclosure of privileged

¹⁴³ Mazza & Sternberg, *supra* note 84.

¹⁴⁴ See, e.g., *Wiginton v. C.B. Richard Ellis*, 229 F.R.D. 568, 570-72 (N.D. Ill. 2004) (noting the low responsive rate in a sexual harassment case where parties used agreed-upon keyword list including “sex,” “kiss,” “breast,” “porn,” “behavior,” “discipline,” “inappropriate,” and “oral”).

¹⁴⁵ Isom, *supra* note 36, at 11.

information.¹⁴⁶ Instead of entering into potentially risky “quick peek” agreements and in addition to “clawback” agreements,¹⁴⁷ parties could use concept search technology to supplement attorney review to guard against the inadvertent disclosure of privileged and confidential documents. This would be especially appropriate in a large matter, where all of the confidential terms, such as attorneys’ names, are not known, and therefore could be missed using a keyword search.

3. DEFENDING THE USE OF CONCEPT SEARCH TECHNOLOGY

[63] There is a perception that discovery of electronic information is the “wild, wild west”¹⁴⁸ of modern litigation. This largely may be due to the fact that the development of technologies used in the discovery of electronic information far outpaces the developing law of discovery.¹⁴⁹ As a result, many of the cutting-edge technologies available to meet the demands of modern discovery have no judicial decisions to support their employment. Litigants are therefore understandably wary of using them.

[64] Unlike keyword search technology, concept searching has not yet been vetted by the courts. While keyword searching is familiar to judges, “it’s very difficult to explain to a judge that a lawyer has searched thoroughly when it’s difficult to explain how concept search technology

¹⁴⁶ To some extent, the recent changes to the Federal Rules of Civil Procedure reflect these new realities by enacting provisions that promote flexibility among parties and increased protection with respect to privilege. For example, amendments to Rule 16(b)(6) contemplate an agreement for waiving privilege in a scheduling order. Likewise, new Rule 26(f) makes specific reference to discovery planning around privilege, as does Rule 26(b).

¹⁴⁷ See THE SEDONA PRINCIPLES, *supra* note 5, cmt. 10.d. (“In a ‘clawback’ (or ‘quick peek’) production, documents are produced to the opposing party before or without a review for privilege, confidentiality, or privacy. The key component of such a production is the ‘clawback’ agreement, in which the parties set stringent guidelines and restrictions to prevent the waiver of confidentiality and privilege. The assumption of the parties to such a ‘clawback’ agreement is that if the requesting party finds a document that appears to be privileged, the producing party can “claw back” the document without having waived any privilege.”).

¹⁴⁸ See Ross Chafin, *The Growth of Cost-Shifting in Response to the Rising Cost and Importance of Computerized Data in Litigation*, 59 OKLA. L. REV. 115, 115 (2006).

¹⁴⁹ Isom, *supra* note 36, at 10.

works.”¹⁵⁰ Nevertheless, courts do recognize that there are various ways to manage electronic documents, and thus many ways in which a party may comply with its obligations.¹⁵¹ Parties are expected to conduct reasonable and adequate searches of their data. So long as a party (or an independent third party vendor) can demonstrate the reasonableness of the approach employed for a given litigation, a party can defend its use of a particular search technology. As concept searching technology becomes standardized, and increasingly familiar to litigants, one may expect that courts will demonstrate a willingness to adopt concept searching as part of the discovery process.¹⁵²

[65] This issue does, however, signal that where possible, a party should attempt to discuss or negotiate appropriate search methodologies, including concept searching, with the opposing party. Courts are far less likely to question the propriety of a given search technology if both parties are in agreement that it is reasonable to expect the search methods used will retrieve relevant information. To the extent parties in litigation determine that the use of concept search technology is a mutually agreeable means for achieving both the producing and requesting parties’ goals of increasing speed and decreasing the cost of document production, they also should discuss whether the cost of that technology should be split rather than borne solely by the responding party.

[66] Moreover, critical to understanding and defending the use of cost-cutting search strategies is the recognition that it is impossible to frame and execute *any* search strategy—human or technological—with perfection. There is already a pronounced level of unreliability and guesswork involved in a traditional full-scale document review process. For example, sorting and reviewing documents by custodian and date can be very problematic: from the reviewer’s perspective, the documents appear randomly and often out of context, making it difficult to

¹⁵⁰ Jason Krause, *Grasping the Concept: The Best Tool for Discovery is Still a Good Lawyer’s Brain*, 91 A.B.A.J. 59, 59 (July, 2005).

¹⁵¹ See *Zubulake III*, 216 F.R.D. 280, 290 (S.D.N.Y. 2003); see also THE SEDONA PRINCIPLES, *supra* note 5, cmt. 6.a.

¹⁵² See Bennett, *supra* note 42, at 16.

consistently spot relevant documents.¹⁵³ Likewise, it is difficult to ensure consistency across the board where a bevy of reviewers, all with different thought processes and judgments as to what may be relevant, are involved in the document review.

[67] Technology will not solve these reliability concerns, but its goal is to improve upon them and reduce the costs of discovery at the same time.¹⁵⁴ Indeed, keyword and concept searching may make locating responsive documents even more reliable insofar as the search results can organize documents by their content.¹⁵⁵ Concept-based review of documents arguably allows for a speedier review with increased chances of spotting relevant documents in context with one another and the document set as a whole, thereby furthering the directives of FRCP 1. Ultimately, “the technology may become so cheap and so ubiquitous that litigants may demand” that keyword and concept searching “be adopted as an essential part of most-e-discovery.”¹⁵⁶

[68] But litigants looking for some kind of “holy grail” in search technology will not find it.¹⁵⁷ There is not, to date, any gold standard for reliability in the discovery of electronic information. Even vendors of search technology software are quick to point out that search technology is not a solution in and of itself, and it only goes so far as the imagination

¹⁵³ Telephone Interview with contributing author to Sedona Conference’s Working Group on Methods for Search and Retrieval (Dec. 8, 2006).

¹⁵⁴ Some studies comparing how human review measures up against automated search techniques suggest that lawyers using only themselves as reviewers fail to do as well as automated techniques in finding relevant documents. Baron, *supra* note 111, at 243. *But see* Telephone Interview with Discovery Vendor (Nov. 27, 2006) (noting that other studies indicate that even the most advanced tools may miss a large percentage of documents that a human would have assessed to be relevant. Another cost-cutting review option offered by vendors is a systemic automated document review that may eliminate a “first-level review” by attorneys for responsive documents. This kind of review protocol focuses on some of the challenges of search technology, namely, consistency and quality assurance, both human nature challenges. Comprehensive studies are underway to test the hypothesis that automated search and retrieval tools can “meet or beat existing human search and retrieval techniques.”).

¹⁵⁵ Baron, *supra* note 111 at 243.

¹⁵⁶ Bennett, *supra* note 42, at 16.

¹⁵⁷ Telephone Interview with contributing author to Sedona Conference’s Working Group on Methods for Search and Retrieval (Dec. 8, 2006).

and adeptness of the lawyer using it.¹⁵⁸ As “with any technology, it is imperative to perform frequent, thorough checks to make sure that the searches are working,”¹⁵⁹ perhaps by using a sampling method, as suggested below. The needs of the litigation at issue should dictate what technology gets used, and how, in order to strike the optimal balance possible between recall and precision.¹⁶⁰ The key to defensibility is that litigants deploy these search strategies as part of a reasonable, good-faith, well-documented discovery protocol. Lawyers must understand where the search technology fits into that protocol and have confidence that they have taken measures to ensure the quality of their searches.

C. SAMPLING

1. USE OF “SAMPLING” IN DISCOVERY TODAY

[69] Litigants have increasingly turned to “sampling” methodologies “to narrow the burden of searching voluminous electronic data for relevant information.”¹⁶¹ Sampling typically involves analyzing a small subset, or sample, of a large amount of ESI, to assess the extent to which relevant documents likely would be found within the remainder of the data. Like concept and keyword searching, sampling is a method to help identify what relevant data may reside in a given population, thereby reducing the potential costs of ESI discovery. Unlike search technologies, however, the goal of sampling is not solely to find potentially relevant data in a way that lives up to performance measurement standards, but rather to make a qualitative judgment as to whether it is worthwhile to conduct further searches of stored electronic data.¹⁶² “By reviewing an appropriate sample of a large body of electronic information, litigants can often determine the

¹⁵⁸ *Id.*

¹⁵⁹ Jason Krause, *Preventing E-Glitches: Understanding Search-Term Basics Ensures More Thorough E-Discovery Compliance*, 92 A.B.A. J. 24, 24 (April, 2006).

¹⁶⁰ Telephone Interview with contributing author to Sedona Conference’s Working Group on Methods for Search and Retrieval (Dec. 8, 2006); Telephone Interview with Discovery Vendor (Nov. 27, 2006).

¹⁶¹ THE SEDONA PRINCIPLES, *supra* note 5, cmt. 11.b.

¹⁶² *See, e.g.,* McPeck v. Ashcroft, 202 F.R.D. 31, 34 (D.D.C. 2001).

likelihood that a more comprehensive review of the materials will yield useful information.”¹⁶³

[70] Sampling methodologies increasingly have been accepted by courts as an appropriate component of a strategy for determining the scope of potentially relevant ESI in computer network backup tapes.¹⁶⁴ Sampling can greatly reduce discovery costs by determining which backup tapes are likely to contain relevant documents, thereby potentially eliminating the costly process of restoring large numbers of tapes and reviewing massive amounts of data. Sampling further provides parties with an assessment of the costs required to restore and review ESI. Armed with that knowledge, courts can make a judgment first as to whether further discovery is warranted, and second, as to which party should bear the costs of restoration, review, and production.¹⁶⁵ Accordingly, the growing trend is for courts to use sampling methodologies to inform cost-shifting analysis in discovery.¹⁶⁶

2. NOT JUST FOR BACKUP TAPES: EXPANDING THE USE OF SAMPLING

[71] The judicial endorsement of sampling in discovery recognizes that the law should not require parties to spend inordinate sums of money to locate, process, and review massive amounts of hard-to-reach data with no assurance that any meaningful portion of that data is relevant to the

¹⁶³ THE SEDONA PRINCIPLES, *supra* note 6, comment 11.b. *See, e.g., McPeck*, 212 F.R.D. at 35 (declining to order additional searches of backup tapes where a “test run” using the sampling method indicated it was unlikely that additional backup tapes contained relevant data); *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 570 (N.D. Ill. 2004) (determining the utility of processing archived material after employing a sampling method using keyword terms).

¹⁶⁴ *See, e.g., Treppel v. Biovail Corp.*, 233 F.R.D. 363, 374 (S.D.N.Y. 2006); *Hopson v. Mayor & City Council of Baltimore*, 232 F.R.D. 228, 245 (D. Md. 2005).

¹⁶⁵ *See, e.g., Wiginton*, 229 F.R.D. at 574-77 (ordering that requesting party should bear 25%, and producing party 75%, of the costs associated with restoring, searching, and managing data on backup tapes after the sampling method revealed costs associated with restoring backup tapes and pertinent documents contained thereon); *Zubulake IV*, 217 F.R.D. 309, 324 (S.D.N.Y. 2003) (“Requiring the responding party to restore and produce responsive documents from a small sample of backup tapes will inform the cost-shifting analysis”).

¹⁶⁶ *See, e.g., Hagemeyer N. Am., Inc. v. Gateway Data Sci. Corp.*, 222 F.R.D. 594, 601-03 (E.D. Wis. 2004). *See infra* Section IV.

lawsuit. Yet most of the case law approving sampling arises in the context of determining whether, and at which party's cost, backup tapes should be searched for relevant ESI.¹⁶⁷ That limitation fails to appreciate that sampling is not just for backup tapes—it can and should play a role in a number of the stages of discovery:¹⁶⁸

[72] *Document Preservation*: Sampling can be used to assist litigants in defining the scope of their document preservation obligation. Sedona Principle 11 explicitly endorses the use of data sampling as a reasonable method by which a litigant may satisfy its good faith obligation to preserve potentially relevant documents.¹⁶⁹ The comments to Principle 11 explain how, for example, sampling may “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.”¹⁷⁰ Understanding whether relevant data resides on voluminous, stored electronic data promotes cost-efficient, intelligent decision-making. This serves the mandates of both FRCP 1 and the new FRCP Amendments.

[73] *Framing Document Production*: Sampling also has potential as a powerful negotiating tool for parties that disagree about the scope of a reasonable discovery search or even regarding the scope of production.¹⁷¹ Rather than search through entire catalogues of extensive ESI for all

¹⁶⁷ See, e.g., *McPeck*, 212 F.R.D. at 37; *Delta Fin. Corp. v. KPMG LLP*, 819 N.Y.S.2d 908, 917 (N.Y. Sup. Ct. 2006).

¹⁶⁸ See Herbert L. Roitblat, *Electronic Discovery Pragmatics Under the New Rules*, 5 (2006) (on file with authors).

¹⁶⁹ See THE SEDONA PRINCIPLES, *supra* note 5, princ. 11.

¹⁷⁰ *Id.* at cmt. 11.a. See also *Hussey v. Chase Manhattan Bank*, No. Civ. A. 02-7099, 2004 WL 220845, *3 (E.D. Pa. Jan. 12, 2004) (holding that it was unnecessary to search an electronic email archive when an identical paper archive had already been searched).

¹⁷¹ See, e.g., *Semsroth v. City of Wichita*, No. 04-1245-MLB-DWB, 2006 WL 3913444, at *11 n.21 (D. Kan. Nov. 15, 2006) (noting that in a dispute over a search of email on backup tapes, the parties could have agreed to run a search of a small sample of e-mails to assist the parties and the court in determining the likelihood that additional searches would disclose important relevant information); *Thompson v. Jiffy Lube Int'l*, No. 05-1203-WEB, 2006 WL 1174040 *6, (D. Kan. May 1, 2006) (ordering parties to confer on whether sampling of electronic database would suffice as method to discern the relevancy of data contained therein).

relevant data, a producing party could offer a sample of its collection and place the burden on the requesting party to show good cause for further production.¹⁷² For example, a party could agree to produce relevant data from a few key custodians as an initial sample, on the basis that any subsequent discovery requests issued would be anchored to what was learned from the initial production. In this vein, sampling could “short-circuit the use of broad requests, multiple rounds of discovery, or multiple depositions that would otherwise be needed to identify the truly responsive documents.”¹⁷³ These uses of sampling would help to ensure that reasonably complete responses are provided to discovery requests while reducing the burden and expense of discovery.¹⁷⁴

[74] *Quality Assurance*: Sampling also may be used to “test the effectiveness of the selection strategies used to identify documents to be reviewed.”¹⁷⁵ For example, a litigant could sample a portion of a collection of electronic documents, review those documents for relevance, then run its keyword or concept search in the same sample to compare the results. Such a measure, taken proactively, could go a long way in both demonstrating the legitimacy of the search technology employed and showing that the party is using a good-faith, reasonable approach to discovery protocols in general.

[75] In whatever discovery context sampling is employed, it is important that litigants use it with diligence and document the protocols used. This will help to defend against any future claims that the sampling was specifically tailored to skew the results.¹⁷⁶

¹⁷² *Gonzales v. Google, Inc.*, 234 F.R.D. 674, 682 (N.D. Cal. 2006) (granting a motion to compel production of a sample of electronic data to one requesting party to assess relevance to the issues of the case).

¹⁷³ Roitblat, *supra* note 168, at 5.

¹⁷⁴ *Id.*

¹⁷⁵ *Id.*

¹⁷⁶ *See* *Quinby v. WestLB AG*, 04 Civ. 7406, 2006 WL 2597900, at *12-13 (S.D.N.Y. Sept. 5, 2006) (noting an argument that a party sample was inherently flawed where the period of sampling was not representative of timeframes covered by the requested document production).

3. A “HYBRID” APPROACH: BLENDING SAMPLING WITH SEARCH TECHNOLOGY TO MAXIMIZE EFFICIENCY AND ACCURACY

[76] The decision of whether to use keywords, concepts, or sampling in a given litigation should not be approached from a “one size fits all” perspective. “[T]here is nothing to prevent combining one or more . . . techniques in hybrid fashion to optimize search results”¹⁷⁷ during both document preservation and review. As two observers put it:

With large text-based data compilations, word and concept searches are well-accepted methods for locating and retrieving responsive data. To narrow large collections of backup media, sampling is likewise a well-accepted methodology. Technology is rapidly developing in this area, and promises to give parties better tools to reduce the scope of searches, reducing costs and burdens all around.¹⁷⁸

[77] Accordingly, courts and litigants are beginning to fashion discovery search protocols for attorney review and production that incorporate both sampling and search technology.¹⁷⁹ For example, due to the enormous costs often associated with data restoration and attorney review, a company can restore a sample of ESI and then use keyword searching to determine what measure of relevant data that sample contains.¹⁸⁰ The company can then extrapolate from that the potential responsiveness of the ESI, as well as how much it would cost an attorney to review the culled-down dataset.¹⁸¹

[78] A combination of sampling and keyword/concept searching also could be used to validate proposed keywords or concepts to be used in culling down a large collection to a smaller set to be reviewed by attorneys. For example, a producing party could apply a certain set of

¹⁷⁷ Baron, *supra* note 111, at 5.

¹⁷⁸ Carroll, *supra* note 102, at 8.

¹⁷⁹ *See, e.g.*, *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 570 (N.D. Ill. 2004) (using a keyword list to locate relevant documents in a sampling of backup tapes).

¹⁸⁰ *See id.*

¹⁸¹ *See id.*; *see also Hagemeyer N. Am., Inc. v. Gateway Data Sci. Corp.*, 222 F.R.D. 594, 603 (E.D. Wis. 2004).

keywords and/or concepts to cull down a sample of the collection and then analyze the results. To the extent the results show that the use of those keywords and concepts did not except a large volume of relevant information from the attorney review process, that validates and supports the use of those keywords and concepts to cull down the remainder of the collection.

[79] Likewise, in the preservation context, companies that have certain technological capabilities can make innovative use of keyword searches, concept searches, and sampling as part of the “reasonable steps” taken to ensure that potentially responsive ESI is located and preserved.¹⁸² Again, the key to reasonableness and defensibility is not necessarily the type of search technology used, but rather the development of a well thought-out, comprehensive search protocol that is applied consistently by the lawyers and companies involved.¹⁸³

[80] The precise combination of methodologies used for a given litigation will depend upon the nature of the case and the data systems of the particular company or companies involved. As Sedona Principle 6 recognizes, it is the producing party who is “best situated to evaluate the procedures, methodologies and technologies appropriate for preserving and producing their own electronic data and documents.”¹⁸⁴ Accordingly, producing parties should strike a balance between a good-faith, cooperative spirit on the one hand, and efforts to avoid allowing the opposing party to dictate the precise methods by which it locates relevant documents on the other.¹⁸⁵

[81] Litigants should educate themselves and courts alike with respect to the potential efficiencies to be gained by employing advanced search strategies in all areas of discovery. Search technologies, while in their infancy, are rapidly growing ever more robust. It is to be expected that the courts will be playing “catch-up” for some time to come. Increasingly, however, “judges, even those raised before the dawn of the modern

¹⁸² See *Zubulake V*, 229 F.R.D. 422, 432 (S.D.N.Y. 2004).

¹⁸³ See *id.*

¹⁸⁴ THE SEDONA PRINCIPLES, *supra* note 5, princ. 6.

¹⁸⁵ Carroll, *supra* note 102, at 3.

computer era, are becoming comfortable with the technology and size of electronic discovery.”¹⁸⁶ Both parties and courts should continue to encourage and promote the use of search and sampling technologies in a variety of stages of the discovery process.¹⁸⁷

D. STRATEGIES TO SUPPORT THE DEFENSIBILITY OF USING SEARCH AND SAMPLING METHODOLOGIES IN DISCOVERY

1. ACT EARLY, ACT COOPERATIVELY

[82] As early as possible, a producing party should consider engaging the requesting party in discussions regarding specific search and sampling methodologies to be used in the discovery of ESI.¹⁸⁸ In addition to the fact that the FRCP Amendments encourage early discussions of this nature,¹⁸⁹ proactively suggesting search and sampling protocols places the

¹⁸⁶ Steven C. Bennett & Thomas N. Niccum, *Two Views from the Data Mountain*, 36 CREIGHTON L. REV. 607, 619 (2003). *See, e.g.*, *Zakre v. Norddeutsche Landesbank Girozentrale*, No. 03 Civ. 0257, 2004 WL 764895 (S.D.N.Y. Apr. 9, 2004).

¹⁸⁷ *See* THE SEDONA PRINCIPLES, *supra* note 6, cmt. 11.a.

¹⁸⁸ Commentators encourage collaborating about search selection criteria with opposing counsel during early “meet and confer” conferences. *See, e.g.*, Gregory P. Joseph & Barry F. McNeil, *Electronic Discovery Standards – Draft Amendments to Civil Discovery Standards*, A.B.A. SECT. LITG. 8 (2003), available at <http://www.abanet.org/litigation/taskforces/electronic/document.pdf> (noting that during early discovery conferences, “parties should consider stipulating to . . . the use of specified key terms or other selection criteria to search some or all of the potentially responsive data for discoverable information.”). Moreover, the FRCP Amendments increase the odds that counsel will be able to settle on detailed search methodologies in the early onset of litigation in light of the need for counsel to become familiar with their clients’ computer systems and ESI for the 26(f) conference, and develop strategies accordingly. *See* FED. R. CIV. P. 26(f) comm. n (“When a case involves discovery of [ESI], the issues to be addressed during the Rule 26(f) conference depend on the nature and extent of the contemplated discovery and of the parties’ information systems. It may be important for the parties to discuss those systems, and accordingly important for counsel to become familiar with those systems before the conference”).

¹⁸⁹ *See* FED. R. CIV. P. 26(f)(3) (requiring the parties to “meet and confer” to discuss any issues relating to disclosure or discovery of electronically stored information); FED. R. CIV. P. 26(f) comm. n (discussing a variety of issues regarding ESI that deserve attention during the discovery planning stage, including identification “of the various sources of [topics relating to discovery] within a party’s control that should be searched for ESI,” and the discussion of “any issues regarding preservation . . . particularly . . . with regard

producing party in a favorable position should future disputes arise regarding ESI.¹⁹⁰ Simply creating an opportunity for the requesting party to have input on those matters enhances the producing party's own credibility in those disputes. The more the requesting party rebuffs attempts at bilateral discussions, the less credibility it will have in making future objections to the court about the search process ultimately selected.¹⁹¹ Parties "who appear to stonewall or fail to make complete responses to discovery requests do so at their peril. The courts can and do penalize recalcitrant parties."¹⁹²

[83] As standards for using particular software products employing one or more search methodologies becomes streamlined, opposing parties in litigation will "theoretically be more likely to reach agreement . . . concerning the use of those products for conducting wide-scale searches of e-records."¹⁹³ Moreover, if the FRCP Amendments are any indication, courts may expect the parties to agree upon streamlined electronic discovery protocols during early meet-and-confer sessions. After all, one of the driving principles behind the FRCP Amendments was the need to expedite discovery and reduce the costs of litigation.¹⁹⁴

to ESI. The volume and dynamic nature of ESI may complicate preservation obligations The parties should pay particular attention to the balance between the competing needs to preserve relevant evidence and to continue routine operations critical to ongoing activities").

¹⁹⁰ See, e.g., *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 374-75 (S.D.N.Y. 2006) (implementing a search protocol formulated by a producing party and noting that the requesting party's refusal to cooperate would have justified the producing party's unilateral search, review, and production of electronic data where requesting party refused to stipulate to a search strategy for electronic documents).

¹⁹¹ See, e.g., *Tulip Computers Int'l v. Dell Computer Corp.*, No. 00-981-RRM, 2002 WL 818061, at *6-7 (D. Del. Apr. 30, 2002) (ordering extensive discovery in the fashion suggested by the requesting party in light of the producing party's earlier failures to cooperate).

¹⁹² Kenneth K. Dort & George R. Spatz, *Discovery in the Digital Era: Considerations for Corporate Counsel*, 20 COMPUTER & INTERNET LAW. 11, 11 (2003); see, e.g., *Coleman Holdings, Inc. v. Morgan Stanley & Co.*, No. 03-5045 AI, 2005 WL 674885 (Fl. Cir. Ct. Mar. 23, 2005) (entering default judgment against producing party for host of discovery abuses, including repeated unjustified refusals to agree to discovery search protocols).

¹⁹³ Baron, *supra* note 111, at 244.

¹⁹⁴ See FED. R. CIV. P. 26(f) comm. n.

[84] The Rules contemplate that a party should be prepared, early in the discovery process, to discuss proposed protocols for streamlining ESI discovery, including search and sampling methodologies and the format for production.¹⁹⁵ As one court recently stated:

[A]s the [FRCP Amendments] make clear, counsel have a duty to take the initiative in meeting and conferring to plan for appropriate discovery of electronically stored information at the commencement of any case in which electronic records will be sought At a minimum, they should discuss . . . the burdens and expenses that the producing party will face . . . and how they may be reduced (i.e. . . . using sampling to search, rather than searching all records).¹⁹⁶

[85] The same court noted that “[t]he days when the requesting party can expect to ‘get it all’ and the producing party to produce whatever they feel like producing are long gone.”¹⁹⁷ Where a defendant has an immense volume of electronic records and the plaintiff relatively few, “it is incumbent upon the plaintiff to have reasonable expectations as to what should be produced by the defendant” in recognition of the fact that, in such cases, ESI discovery “is not played on a level field.”¹⁹⁸

[86] In view of advancing the efficiency directives of FRCP 1, parties should take care to discuss, in detail, search and sampling protocols relating to ESI in meet-and-confer conferences following the issuance of document requests. To the extent that the parties cannot agree on those protocols, responding parties should not unduly hesitate in asking courts to endorse them in furtherance of the spirit and mandate of FRCP 1. Raising and resolving these issues at the front-end of discovery not only reduces the costs of managing relevant data, but it also may serve to diffuse potential discovery disputes further down the road, thereby advancing the

¹⁹⁵ *Id. See, e.g.,* Hopson v. Mayor & City Council of Baltimore, 232 F.R.D. 228, 245 (D. Md. 2005).

¹⁹⁶ *Hopson*, 232 F.R.D. at 245.

¹⁹⁷ *Id.*

¹⁹⁸ *Id.*

directive of FRCP 1 to “secure the just, speedy, and inexpensive determination of every action.”¹⁹⁹

2. DO NOT WAIT UNTIL THE MOTION TO COMPEL: TAKE EARLY,
AFFIRMATIVE STEPS TO SECURE THE COURT’S ENDORSEMENT OF SEARCH
AND SAMPLING STRATEGIES

[87] Any seasoned litigator knows that discovery negotiations do not take place in Shangri-La. It is not difficult to imagine a scenario where, despite numerous meet-and-confer sessions, the parties cannot arrive at an agreed-upon protocol for the use of search and sampling methodologies for the discovery of ESI. While parties can always enter into court-endorsed stipulations regarding electronic discovery protocols after the Rule 26(f) and Rule 16(b) timeframe,²⁰⁰ parties often wait until they are in a motion to compel context before they get the court involved. There is no reason to so delay court participation, especially now in light of the recent FRCP Amendments. Absent agreement from the opposing party, parties should start acting affirmatively to secure court approval of reasonable discovery protocols, including search strategies involving sampling, keyword, and/or concept search methodologies.²⁰¹

[88] For example, a party can move for a protective order²⁰² outlining specific terms and conditions under which discovery of electronic information may be had. New Rule 26(b)(2)(B) explicitly provides that a party from whom discovery is sought may move for a protective order from discovery of ESI from sources that it identifies as “not reasonably accessible because of undue burden or cost.”²⁰³ Even where ESI is “accessible,” but nonetheless costly to search, review, and produce, a party may always move for a protective order under the proportionality rule, now found in Rule 26(b)(2)(C).²⁰⁴ The proportionality rule recognizes

¹⁹⁹ FED. R. CIV. P. 1.

²⁰⁰ *See, e.g.*, Palgut v. City of Colo. Springs, No. CIVA 06CV01142 WDMMJ, 2006 WL 3483442, at *2 (D. Colo. Nov. 29, 2006) (discussing a stipulation and order regarding electronic discovery plan and order to preserve evidence).

²⁰¹ *See* Treppel v. Biovail Corp., 233 F.R.D. 363, 374 (S.D.N.Y. 2006).

²⁰² FED. R. CIV. P. 26(c).

²⁰³ FED. R. CIV. P. 26(b)(2)(B).

²⁰⁴ FED. R. CIV. P. 26(b)(2)(c).

that discovery should be limited where the court determines that burden or expense of discovery outweighs its likely benefit.²⁰⁵ A motion for protective order should detail the proposed search procedures, explain why attempts to secure agreement from the opposing party failed, provide specifics as to the costs and burdens associated with the proposal, and otherwise explain good cause for the terms proposed.

[89] As discussed in section IV.C, *infra*, the proportionality rule set forth in 26(b)(2)(C) and 26(c) also allows a court to deny altogether an aspect of discovery where the burdens of the discovery outweigh its benefits.²⁰⁶ Obtaining the court's outright denial of a form or scope of discovery requested by the opposing party may in some cases be the most appropriate way to achieve the "just, speedy, and "inexpensive" resolution mandated by FRCP 1.

E. ATTAINING "DIGITAL ENLIGHTENMENT"²⁰⁷ WITH AN EFFECTIVE DATA LIFECYCLE MANAGEMENT POLICY

[90] Perhaps the most effective measure a company can take to cut discovery costs has nothing directly to do with discovery, but with the general day-to-day management of its data lifecycle. As discussed above, modern technologies have done much to facilitate business, but "these same technologies have the capacity to create a nightmarish amount of potentially discoverable material."²⁰⁸ In-house counsel can save a company a lot of money (and headaches) by developing and implementing a robust data lifecycle management policy/program that provides guidelines and requirements for the creation, dissemination, storage, archival, and disposal of the company's ESI. "Such policies and procedures, if followed, reduce the amount of stored electronic information and streamline the production process in the event of litigation."²⁰⁹

²⁰⁵ *Id.*

²⁰⁶ See *infra* section IV.C.4.

²⁰⁷ Ashby Jones, *What a Mess!*, NAT'L L. J., Dec. 2, 2002, at C6.

²⁰⁸ Dort & Spatz, *supra* note 192, at 15.

²⁰⁹ *Id.* at 13.

[91] The FRCP Amendments have generated some panic and confusion among companies, and their in-house counsel, about what changes companies are now required to make with respect to data lifecycle management. Will companies now need to retain, in perpetuity, all e-mails and other ESI generated by their employees (as one news report incorrectly implied the day the new Rules went into effect)?²¹⁰ The answer, assuredly, is “no.” Of course, and as discussed above, as soon as a company reasonably anticipates litigation it *must* institute a “litigation hold” against any aspect of its data lifecycle management program that would result in the spoliation of potentially relevant ESI.²¹¹ The Amendments do not change that obligation. But there is no Rule (Amended or otherwise) that dictates how companies must manage data outside of litigation.

[92] Though the Rules do not require a company to implement a data lifecycle management program, it is fair to say that the Amendments (particularly new Rule 37(f)) make it even more beneficial for a company to do so. Today companies “sit atop volcanoes of bits and bytes.”²¹² Companies use and retain ESI in a variety of different forms, and in a multiplicity of locations, including servers, backup tapes, desktops, laptops, PDAs, home computers, and so forth, many of which are often hard to reach.²¹³ Data storage is relatively inexpensive, so it often “just piles up,” turning “corporate headquarters into technology silos.”²¹⁴ Such hodgepodge document management is increasingly less workable in light of the modern liabilities associated with the discovery of electronic information. When ESI is managed poorly, it can end up later costing a company millions of dollars in excess discovery costs when litigation hits.

²¹⁰ See, e.g., *Careful: Employee E-Mails, IMs Must Be Tracked*, Dec. 1, 2006, <http://listserv.educause.edu/cgi-bin/wa.exe?A2=ind06&L=smallcol&D=1&P=38497> (last visited March 19, 2007) (warning falsely, that “U.S. companies will need to keep track of all of the e-mails, instant messages and other electronic documents generated by their employees thanks to [the amended FRCP that took effect December 1, 2006]”).

²¹¹ See *supra* notes 83 & 85 and accompanying text.

²¹² Jones, *supra* note 207.

²¹³ Mazza & Sternberg, *supra* note 84.

²¹⁴ Jones, *supra* note 207.

[93] First, the lack of an effective data lifecycle management program makes the task of finding and preserving data potentially relevant to a specific litigation like finding a needle in a haystack. Many of the high costs associated with discovery arise in part because electronic data stored for one purpose is suddenly required to be preserved, collected, and processed for another purpose, and the company infrastructure is not able to shift gears accordingly.²¹⁵ Simply put, “the larger the number of records stored, the slower and more rigorous the search for them will be.”²¹⁶ Thus, companies end up spending a great deal of money simply to locate their own stored information, often without any guarantee of reasonable success.²¹⁷ Moreover, even when companies do have data lifecycle management policies, such as data retention schedules, a lack of oversight and enforcement in making sure employees carry out those policies can result in the “expenditure of considerable time and money in discovery disputes over the production of e-mails that should have been destroyed in the first place.”²¹⁸

[94] Second, plaintiffs’ lawyers are increasingly taking advantage of the apparent mismanagement or non-management of ESI to cause litigation to degenerate into a collateral proceeding about information that the company allegedly has allowed to be spoliated and what kind of “adverse inference” or other sanction is an appropriate punishment for that

²¹⁵ Isom, *supra* note 36, at 11.

²¹⁶ Marilee S. Chan, *Paper Piles to Computer Files: A Federal Approach to Electronic Records Retention and Management*, 44 SANTA CLARA L. REV. 805, 809 (2004).

²¹⁷ Dort & Spatz, *supra* note 192, at 3.

²¹⁸ *Id.* For example, in *Tulip Computers*, defendant Dell belatedly identified the existence of a massive amount of potentially responsive data nine months after being served with discovery requests. Dell admitted that it had mistakenly destroyed documents as well. *Tulip Computers Int’l B.V. v. Dell Computer Corp.*, No. CIV.A. 00-981-RRM, 2002 WL 818061, at *4 (D. Del. April 30, 2002). In granting the plaintiff extensive additional discovery that allowed it “to ascertain for themselves whether Dell’s representations that all responsive documents have been produced are accurate,” noting Dell’s failure “indicates either a failure to take its discovery obligations with the required degree of seriousness and diligence or an extreme lack of knowledge and control over its own files and procedures.” *Id.* at *6-7; see also *Murphy Oil USA, Inc. v. Fluor Daniel, Inc.*, No. Civ.A. 99-3564, 2002 WL 246439, *2 (E.D. La. Feb. 19, 2002) (stating that expensive and time-consuming discovery dispute over e-mails “would be moot” if producing party had followed its e-mail retention policy).

spoliation.²¹⁹ In the context of sanctions for the spoliation of relevant documents, a company's ignorance about the way computer systems retain and dispose of ESI is not a defense. The seminal *Zubulake* decisions put an end to any doubts about that.²²⁰ Irrespective of the state of a company's management of records, once it "reasonably anticipates litigation, it *must* suspend its routine document retention/destruction policy and place a 'litigation hold' to ensure the preservation of relevant documents."²²¹ This duty requires "a party and her counsel [to] make certain that all sources of potentially relevant information are identified and placed 'on hold.' To do this, counsel must become fully familiar with her client's document retention policies, as well as the client's data retention architecture."²²²

[95] A key problem with this standard, for many companies, is that they often have no coherent data retention policy to begin with—in essence, they do not have their ESI "house" in order. In a 2006 survey published by E-Discovery Advisor Magazine, 50% of corporate attorneys responded that they are either not confident in their company's document retention program or their company doesn't have any program/policy.²²³ Compounding the problem, many in-house lawyers "are very uncomfortable with the technical side" of data management.²²⁴

[96] It is vital that corporate America begin to regain control of its electronic data.²²⁵ The best measure companies can take to avoid being trampled underfoot by both mounting discovery costs and the increasing risk of sanctions for spoliation of relevant evidence is to develop and implement an effective data lifecycle management program that addresses the systemic storage and disposal of electronic data. While a detailed discussion of the elements of data lifecycle management are beyond the

²¹⁹ Mazza & Sternberg, *supra* note 84, at 19.

²²⁰ *Zubulake IV*, 220 F.R.D. 212 (S.D.N.Y. 2003); *Zubulake V*, 229 F.R.D. 422 (S.D.N.Y. 2004).

²²¹ *Zubulake IV*, 220 F.R.D. at 218 (emphasis added).

²²² *Zubulake V*, 229 F.R.D. at 432.

²²³ *Are Companies Prepared for the New Federal Rules on E-Discovery?*, E-DISCOVERY ADVISOR MAG., Web Ed. 2006, Week 48, <http://e-discoveryadvisor.com/doc/18626>.

²²⁴ Jones, *supra* note 207, at C7.

²²⁵ *Id.* at C6.

scope of this article,²²⁶ the development and implementation of a successful program will require participation from multiple departments within a company, including legal, IT, and senior management. “All who are involved in the creation, storage, and destruction of documents should follow the policy, and the company should create an enforcement mechanism.”²²⁷ It is therefore important for companies to take affirmative steps to confirm that their policies are understood and followed by employees at all levels.²²⁸ IN the end, a program can keep ESI volume and proliferation to a minimum so that there is less data sitting around, and in fewer locations, when litigation hits, thus serving as the first line of defense against the increasing costs associated with ESI discovery.

IV. COST-SHIFTING IN DISCOVERY: WINNING THE BATTLE THAT MAY WIN THE WAR

[97] A long-standing federal presumption, articulated by the U.S. Supreme Court in *Oppenheimer Fund, Inc. v. Sanders*, holds that the party responding to a discovery requests pays all related production expenses.²²⁹ Modern courts, however, have shown an increased willingness to issue cost-shifting orders to reduce the burden of discovery for parties that otherwise would be saddled with huge discovery costs given the scope of the discovery requests and/or the inaccessibility of the potentially relevant ESI. This trend recognizes that a party should not be able to strategically

²²⁶ See Dort & Spatz, *supra* note 192, at 13 (discussing and listing the considerations corporate counsel should take into account when developing a document retention policy)

²²⁷ Dort & Spatz, *supra* note 192, at 13.

²²⁸ Jones, *supra* note 207, at C7 (“[I]t takes just one provocative e-mail to create a public-relations disaster or a litigation liability.”). See *Zubulake V*, 229 F.R.D. at 439-40 (discussing the liabilities caused when employees disregard company directives in an atmosphere where the directives are not properly monitored and enforced by the company).

²²⁹ See *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 358 (1978) (holding under discovery rules that “the presumption is that the responding party must bear the expense of complying with discovery requests”); *Toshiba Am. Electronic Components, Inc. v. Super. Ct.*, 21 Cal. Rptr. 3d 532, 538 (2004) (“The general rule in both state and federal court is that the responding party bears the expense typically involved in responding to discovery requests, such as the expense of producing documents.”).

abuse the *Oppenheimer* presumption by issuing broad discovery requests to drive up opponents' litigation costs and force settlements.²³⁰

[98] Cost-shifting battles are hotly contested and for good reason: decisions on motions regarding who will be required to pay for discovery responses (the cost of which may run into the hundreds of thousands, if not tens of millions, of dollars) can impact severely how an action proceeds and in fact may be outcome-determinative in some cases.²³¹ This section examines the evolution of the law of cost-shifting both in federal and state court and as relevant in the recent FRCP Amendments. It then explores the arguments in favor of shifting other ESI discovery costs that have not historically been the subject of cost-shifting discussions. Finally, this section sets out several issues practitioners should consider when making the move to shift all or some of their clients' discovery costs.

A. THE EVOLUTION OF THE LAW OF COST-SHIFTING

[99] "Early" cases addressing cost allocation of ESI discovery required the responding party to bear the cost of producing ESI, reasoning that having to produce the data in litigation was an ordinary and foreseeable risk of using electronic storage media.²³² Despite an increasing number of high-

²³⁰ Maurice Rosenberg, *Federal Rules of Civil Procedure in Action: Assessing Their Impact*, 137 U. PA. L. REV. 2197, 2204 (1989).

²³¹ See, e.g., *Zubulake I*, 217 F.R.D. 309, 317-18 (S.D.N.Y. 2003)

Courts must remember that cost-shifting may effectively end discovery, especially when private parties are engaged in litigation with large corporations. As large companies increasingly move to entirely paper-free environments, the frequent use of cost-shifting will have the effect of crippling discovery in discrimination and retaliation cases. This will both undermine the 'strong public policy favor[ing] resolving disputes on their merits,' and may ultimately deter the filing of potentially meritorious claims.

Id. (quoting *Pecarsky v. Galaxiworld.com, Inc.*, 249 F.3d 167, 172 (2d Cir. 2001)); *Analog Devices, Inc. v. Michalski*, No. 01 CVS 10614, 2006 WL 3287382, *5 (N.C. Sup. Ct. Nov. 1, 2006) ("It is indisputable that the decisions concerning the costs of e-discovery in some cases could be outcome-determinative.").

²³² See, e.g., *In re Brand Name Prescription Drugs Antitrust Litig.*, No. 94 C 897, MDL 997, 1995 WL 376682 (N.D. Ill. Jun. 15, 1995) (denying a motion that requesting party bear the cost of producing electronic documents: "if a party chooses an electronic storage method, the necessity for a retrieval program or method is an ordinary and foreseeable

profile cases and the recent FRCP Amendments turning the tide, the law on shifting the cost of producing ESI remains jurisdiction-specific, often unsettled, sometimes conflicting, and continually evolving. While cost-shifting principles emerging in federal court can provide some guidance for resolving ESI discovery disputes in state courts, state-specific statutes and case law can vary greatly from federal standards. Counsel for a party wishing to effectively pursue the shifting of costs associated with ESI discovery, therefore, should gain an understanding of the emerging standards and recent judicial decisions in the relevant jurisdictions.

1. FEDERAL LAW

[100] The majority of ink that has been devoted to discussing the trials and tribulations of ESI discovery (not to mention the seminars, forums, working groups, rules, guidelines, and opinions) focuses on federal courts and the FRCP. Existing federal case law on the topic of cost-shifting deals almost exclusively with the burden and expense of producing data stored on back-up tapes.²³³ Interestingly, the Rules do not expressly contemplate

risk”); *Bills v. Kennecott Corp.*, 108 F.R.D. 459, 464 (D. Utah 1985) (denying a motion to shift costs of discovery of electronic data because: (1) the defendant was in the “most economical position to call up its own computer stored data,” (2) cost was not excessive, (3) relative burden in obtaining the data was substantially greater to requesting party, and (4) responding party was benefited to some degree); *Linnen v. A.H. Robins Co.*, No. 97-2307, 1999 WL 462015, at *6 (Sup. Ct. Mass. June 16, 1999) (stating that the cost of restoring electronic data in response to discovery request is “one of the risks taken on by companies which have made the decision to avail themselves of the computer technology”).

²³³ See, e.g., *Quinby v. WestLB AG*, No. 04CIV.7406 (WHP) (HBP), 2006 WL 2597900, at *16 (S.D.N.Y. Sept. 5, 2006) (shifting 30% of costs of restoring and searching e-mails of former employees from back-up tapes); *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 574-77 (N.D. Ill. 2004) (shifting part of the costs associated with restoring, searching, and managing data on backup tapes to requesting party); *Hagemeyer N. Am., Inc. v. Gateway Data Sci. Corp.*, 222 F.R.D. 594, 603 (E.D. Wis. 2004) (ordering defendant to restore a sample of backup tapes and parties to make additional submissions addressing whether burden or expense of satisfying entire request is proportionate to likely benefit); *Medtronic Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550, 558 (W.D. Tenn. 2003) (requiring defendant to bear part of costs of producing approximately 996 network backup tapes, containing, among other things, plaintiff’s email); *Zubulake I*, 217 F.R.D. at 322 (articulating seven factor cost-shifting analysis for determining who should pay the cost of producing material from “inaccessible” media, like backup tapes); *Rowe Entm’t, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 423 (S.D.N.Y. 2002)

cost-shifting. Rather, in federal court, discovery cost-shifting has historically been done pursuant to a district court's discretion under Rule 26(c) "to grant orders protecting [respondents] from 'undue burden or expense,'" which has been interpreted to include orders for the requesting party's payment of the costs of discovery.²³⁴

[101] Rule 26(c) protective orders are especially important in the context of ESI discovery, where the costs of production can cripple responding parties and force them to settle cases prematurely. There is a growing body of federal case law addressing cost-shifting when ESI is inaccessible, such as data that resides on back-up tapes. Most notable are the watershed *Zubulake* decisions, which are widely regarded as the most thorough treatment of cost-shifting under federal law. In light of the Amendments addressing cost-shifting considerations under the FRCP, however, the validity of pre-Amendment case law is unclear. Given that *Zubulake's* Judge Scheindlin sat on the committee to amend the Rules, the cases and its progeny still may be valuable tools for interpreting the meaning and scope of the Amended Rules.

(A) ZUBULAKE'S APPROACH

[102] In the oft-cited decisions in *Zubulake v. UBS Warburg*, which are six in number, Judge Scheindlin of the Southern District of New York created the cost-shifting test for discovery that has been widely used by the federal courts.²³⁵

(recognizing the difficulties created when litigants must comply with discovery requests seeking data contained on backup tapes); *McPeck v. Ashcroft*, 202 F.R.D. 31, 34 (D.D.C. 2001) ("The more likely it is that the backup tape contains information that is relevant to a claim or defense, the fairer it is that the [responding party] search at its own expense").

²³⁴ *Oppenheimer Fund*, 437 U.S. at 358. See also FED. R. CIV. P. 34, advisory comm. notes to 1970 Amendments ("[T]he Courts have ample power under Rule 26(c) to protect respondent against undue burden or expense, either by restricting discovery or requiring that the discovering party pay costs.").

²³⁵ See *Zubulake I*, 217 F.R.D.309; *Zubulake II*, 230 F.R.D. 290 (2003); *Zubulake III*, 216 F.R.D. 280 (2003); *Zubulake IV*, 220 F.R.D. 212 (2003); *Zubulake V*, 229 F.R.D. 422 (2004); *Zubulake VI*, 382 F. Supp. 2d 536 (2005). Because the new test required the producing party to investigate certain issues and report back to the court, Judge

[103] The plaintiff in *Zubulake* contended that “key evidence is located in various e-mails exchanged among UBS employees that now exist only on backup tapes and perhaps other archived media.”²³⁶ UBS argued that it had produced all responsive documents, although admittedly it had never searched for responsive emails on any of its backup tapes.²³⁷ The court promptly rejected UBS’s argument, finding the facts “strongly” suggested that the backup tapes contained responsive emails necessary to *Zubulake*’s case.²³⁸ The court then turned to the issue of cost-shifting. UBS urged that, due to the immense cost of restoring and searching the backup tapes, it should not bear the undue burden or expense of producing e-mails from them.^{FN} In evaluating whether to diverge from the presumption that the responding party pays for discovery costs, the court adopted the following three-step approach.

(I) STEP ONE: DETERMINE ACCESSIBILITY OF DATA

[104] *Zubulake* instructed that the “first question . . . is whether cost-shifting must be considered” at all.²³⁹ The court rejected the proposition that shifting should be considered in all ESI cases, citing both the *Oppenheimer* presumption and the fact that today “virtually all cases” involve ESI.²⁴⁰ Rather, the court held that “cost-shifting should only be considered when electronic discovery imposes an ‘undue burden or expense’ on the responding party.”²⁴¹ “[W]hether production of documents is unduly burdensome or expensive turns primarily on whether it is kept in an accessible or inaccessible format (a distinction that corresponds closely to the expense of production).”²⁴²

Scheidlin did not apply it fully to the facts of her case until *Zubulake III*. For cases adopting *Zubulake*’s approach, see *infra* n.252.

²³⁶ *Zubulake I*, 217 F.R.D. at 311-12.

²³⁷ *Id.* at 317.

²³⁸ *Id.*

^{FN} *Id.*

²³⁹ *Id.*

²⁴⁰ *Id.*

²⁴¹ *Id.* at 318.

²⁴² *Id.*

[105] The court then offered the following hierarchy of the accessibility of ESI, from most accessible to least accessible:

- Active, online data (data used in the very active states of its life, such as when it is being created or received or processed, e.g., hard drives);
- Online data (data stored in a robotic storage device that uses multiple read/write devices to store and retrieve records, e.g. optical disks);
- Offline storage/archives (data stored on removal optical disks or magnetic tape media that can be labeled and stored on a shelf or rack);
- Backup tapes (data stored on tape drives that are not organized for retrieval);
- Erased, fragmented or damaged data (data never meant to be retrieved that can only be accessed after significant processing).²⁴³

[106] Of these categories, the court held the first three were considered accessible and the last two were considered inaccessible.²⁴⁴ This portion of the *Zubulake* test is particularly important because it is the gateway to the cost-shifting analysis. If the requested discovery comes from the first three categories, the *Oppenheimer* presumption controls. If the discovery comes from the latter two categories, a cost-shifting analysis ensues.

(II) STEP TWO: TEST-RUN SAMPLE DATA

[107] The second step instructs that the producing party should test a sample of the data to estimate the cost and likelihood that the requested discovery will be responsive. The court held that the results of that test would be used to determine whether to shift costs.²⁴⁵

²⁴³ *Id.* at 318-19.

²⁴⁴ *Id.* at 319-20.

²⁴⁵ *Id.* at 323-24.

(III) STEP THREE: APPLY THE SEVEN-FACTOR TEST

[108] Armed with the information learned in the test-run, the third and final step involves analyzing whether cost-shifting is appropriate.²⁴⁶ Emphasizing the requirements of Rule 26(b)(2)(iii) (now Rule 26(b)(2)(C)(iii)), the court offered its own seven-factor test, listed from most to least important:

- (1) the extent to which the request is specifically tailored to discover relevant information;
- (2) the availability of such information from other sources;
- (3) the total cost of production, compared to the amount in controversy;
- (4) the total cost of production, compared to the resources available to each party;
- (5) the relative ability of each party to control costs and its incentive to do so;
- (6) the importance of the issues at stake in the litigation; and
- (7) the relative benefits to the parties of obtaining the information.²⁴⁷

[109] The court applied its seven-factor test to the sampling results and found that factors 1-4 cut against cost-shifting, factors 5-6 were neutral, and factor 7 favored cost shifting.²⁴⁸ Although these findings would appear to weigh against cost-shifting, the court held that because continued production may produce valuable information, “some cost shifting is appropriate.”²⁴⁹ The court explained that “the precise allocation is a matter of judgment and fairness rather than a mathematical

²⁴⁶ *Id.* at 324.

²⁴⁷ *Id.* at 321-23.

²⁴⁸ *See* Zubulake III, 216 F.R.D 280, 289 (S.D.N.Y. 2003)

²⁴⁹ *Id.*

consequence of the seven factors”²⁵⁰ The court shifted 25% of the cost to the plaintiff and assigned the other 75% to the defendant.²⁵¹

[110] The primary point of the *Zubulake* decision is that a district court judge has broad discretion in deciding cost-shifting disputes. Nevertheless, the *Zubulake* approach has been applied in several federal cases.²⁵²

(B) AMENDMENTS TO THE FEDERAL RULES OF CIVIL PROCEDURE

[111] The Amended FRCP provide that parties are expected to produce relevant documents and data from locations that are “reasonably accessible,” and notify their opponents of locations that are “not reasonably accessible” but where it is possible that relevant documents may reside.²⁵³ “A party need not provide discovery of electronically stored information from sources that the party identifies as not reasonably accessible because of undue burden or cost.”²⁵⁴ The Committee Note to Rule 26(b)(2)(B) suggests that the required identification “should, to the extent possible, provide enough detail to enable the requesting party to evaluate the burdens and costs of providing the discovery and the likelihood of finding responsive information on the identified sources.” It

²⁵⁰ *Id.*

²⁵¹ *Id.*

²⁵² *See, e.g.*, *Semsroth v. City of Wichita*, No. 04-1245-MLB-DWB, 2006 U.S. Dist. LEXIS 83363, at *19-35 (D. Kan. Nov. 15, 2006); *Quinby v. WestLB AG*, No. 04 Civ. 7406, 2006 WL 2597900, at *7, *11-16 (S.D.N.Y. Sept. 5, 2006); *Xpedior Credit Trust v. Credit Suisse First Boston (USA), Inc.*, 309 F. Supp.2d 459, 461 (S.D.N.Y. 2003); *Multitech. Servs., L.P. v. Verizon*, No. Civ. A 4:02-CV-702-Y, 2004 WL 1553480 (N.D. Tex. July 12, 2004); *Wiginton v. CB Richard Ellis, Inc.*, 229 F.R.D. 568, 570-73 (N.D. Ill. 2004) (modifying *Zubulake* test by adding an eighth factor “that considers the importance of the requested discovery in resolving the issues of the litigation”; *Hagemeyer N. Am., Inc. v. Gateway Data Sci. Corp.*, 222 F.R.D. 594, 602-03 (E.D. Wis. 2004); *OpenTV v. Liberate Techs.*, 219 F.R.D. 474, 476 (N.D. Cal. 2003).

²⁵³ FED. R. CIV. P. 26(b)(2)(B). Even where a party asserts that potentially relevant electronic documents or data reside in a location that is “not reasonably accessible,” that party still may be obligated to preserve the data in that location throughout the life of the litigation. FED. R. CIV. PRO. 26(b)(2)(B) (amended 2006) comm. note

²⁵⁴ FED. R. CIV. P. 26(b)(2)(B).

does not specify when (at what point in the discovery process) that identification should take place.²⁵⁵

[112] If an agreement cannot be reached as to whether, or on what terms, sources identified as “not reasonably accessible” should be searched and discoverable information produced, a requesting party may move to compel production or the responding party can move for a protective order barring the discovery.²⁵⁶ The court may order production if the responding party does not convince the court that the data is “not reasonably accessible.” Even if the responding party does so convince the court, the court still may order production where the requesting party shows “good cause” for the production, “considering the limitations set forth in Rule 26(b)(2)(C).”²⁵⁷ What facts will satisfy a responding party’s burden of proving inaccessibility is not spelled out in the Rule.²⁵⁸ The Advisory Committee notes to Amended Rule 26(b)(2) provide that, once the burden has shifted to the requesting party to show good cause, the seven-factor test from *Zubulake* provides some of the “[a]ppropriate considerations” that may be used to determine whether the burdens and costs of requiring a responding party to search for and produce ESI that is not reasonably accessible can be justified in the circumstances of the case.²⁵⁹

[113] If the court orders production, the requesting party may be required to pay some or all of the cost of accessing the data and converting it into a format that allows the responding party to review and produce it.²⁶⁰ The Rules also contemplate that parties discuss, and even experiment with, sampling as a means to assess the best way to deal with data that is “not reasonably accessible.”²⁶¹ As of the time of the writing of this article, there have been no notable federal decisions construing or interpreting Rule 26(b)(2)(B) as Amended.

²⁵⁵ FED. R. CIV. PRO. 26(b)(2)(B) (amended 2006) comm. n.

²⁵⁶ *Id.*

²⁵⁷ FED. R. CIV. P. 26(b)(2)(B).

²⁵⁸ FED. R. CIV. PRO. 26(b)(2)(B) (amended 2006) comm. n.

²⁵⁹ *Id.*

²⁶⁰ *Id.*

²⁶¹ *Id.*

2. STATE COURT

[114] State court judges are seeing disputes regarding the discovery of electronic information with increasing frequency. When those discovery issues arise in state court, however, they are often governed by state rules and case law that are substantially different from those that guide federal courts. In particular, as is discussed in this section, some states have procedural rules governing the allocation of discovery costs that are quite different from the federal rules.²⁶² Moreover, while there are numerous federal decisions addressing who should bear the burden of costs associated with discovery of electronic information, there is comparatively little precedent in state courts.

(A) CALIFORNIA

[115] In California, California Code of Civil Procedure section 2031.280(b) governs whether some or all of the potentially substantial costs associated with producing ESI in a usable format should be shifted to the requesting party. It provides that “any documents demanded shall either be produced as they are kept in the ordinary course of business, or be organized and labeled to correspond with the categories in the demand” and that, “[i]f necessary, the responding party at the reasonable expense of the demanding party shall, through detection devices, translate any data compilations included in the demand into reasonably usable form.”²⁶³ This cost-shifting provision therefore puts the burden on the requesting party to pay for electronic data translation and compilation as long as the cost is reasonable.

²⁶² *But see* Bank of Am. Corp. v. SR Int’l Bus. Ins. Co., No. 05-CVS-5564, 2006 WL 3093174, at *16 (N.C. Sup. Ct. Nov. 1, 2006) (discussing North Carolina’s Guidelines For State Trial Courts Regarding Discovery Of Electronically-Stored Information regarding reallocation of discovery costs reflects three-tiered test and cost-benefit analysis from *Zubulake* and discussing that “[l]ike *Zubulake*, the Guideline treats cost-shifting as a matter for the judge’s discretion”).

²⁶³ CAL. CODE CIV. PROC. § 2031.280(b). California courts may also order the use of technology, such as CD-ROMs, Internet Web sites, electronic document depositories, Internet depositions, videoconferencing, etc., to aid the discovery process in certain cases. *See* CAL. CODE CIV. PROC. §§ 2017.730-740.

[116] The question of whether the cost-shifting contemplated by section 2031.280(b) is mandatory was answered in the affirmative by the California Court of Appeal in *Toshiba America Electronic Components, Inc. v. Superior Court*. In a rare appellate writ proceeding on a discovery issue, the California Court of Appeal considered the phrase “at the reasonable expense of the demanding party” and whether this cost-shifting was “mandatory” or “merely permits the trial court to shift the cost to the demanding party when the responding party objects.”²⁶⁴ Defendant Toshiba argued that the provision automatically shifted costs for extensive data restoration, estimated between \$1.5 and \$1.9 million.²⁶⁵ Plaintiff Lexar Media countered that *Zubulake* and Rule 26(c)’s “undue burden or expense” standard controls.²⁶⁶ While recognizing that “the general rule in both state and federal court is that the responding party bears the expense typically involved in responding to discovery requests[,]”²⁶⁷ the court relied on the controlling California statute to ultimately side with Toshiba, holding that the demanding party is expected to pay the reasonable costs of necessary electronic data translation pursuant to section 2031.280(b), and noting that *Zubulake* does not control in California.²⁶⁸

[117] The *Toshiba* court is said to have “charted a new course in the cost-shifting debate” in California that “could significantly affect the way litigators in the state approach cost-shifting arguments.”²⁶⁹ The case, however, left several key questions unanswered. For example, what are “reasonable” expenses and “necessary” translations under section 2031.280(b)? Does the rule for back-up tapes apply to other forms of ESI, given that section 2031.280(b) refers broadly to “data compilations”? The court did advise that the bounds of “reasonable” and “necessary” should

²⁶⁴ *Toshiba Am. Elec. Components, Inc. v. Lexar Media*, 124 Cal. App. 4th 762, 767 (App. Ct. 2004).

²⁶⁵ *Id.* at 766.

²⁶⁶ *Id.* at 770.

²⁶⁷ *Id.* at 769.

²⁶⁸ *Id.* at 771-72. The *Toshiba* court noted that while California Code of Civil Procedure section 2031 does not contain a specific procedure for challenging the burden that cost-shifting may place on the demanding party, insofar as a demanding party thinks the costs are unreasonable, that party must avail itself of generally applicable relief procedures, such as a protective order. *Id.* at 773.

²⁶⁹ Linda G. Sharp, *Restoration Drama: The Complexity of Electronic Discovery Requires Practitioners to Master New Litigation Skills*, 28 L.A. LAW. 31 (2005).

be determined by sampling and may result in only partial cost-shifting or none at all; indeed, the demanding party may seek a protective order if it disputes either point.²⁷⁰ While any attempts to extend *Toshiba's* holding to other platforms, such as deleted data on hard drives, are certain to meet with stiff resistance, California cases generally recognize that burdensome discovery requests may require cost-shifting to satisfy “principles of fundamental fairness,” since the requesting party is asking the producing party to do something special that falls beyond the call of routine discovery.²⁷¹

(B) NEW YORK

[118] In stark contrast to the *Oppenheimer* presumption and the *Zubulake* approach for determining when to depart from that presumption, New York courts have held that under the state’s Civil Practice Law and Rules (CPLR), “the party seeking discovery should incur the costs incurred in the production of discovery material.”²⁷² As such, whether it is appropriate to shift the costs of discovery of electronic information is not an issue in New York state court because the presumption at the outset is that the requesting party pays for discovery. “Therefore, the analysis of whether electronic discovery should be permitted in New York is much simpler than it is in the federal courts. The court need only determine whether the material is discoverable and whether the party seeking the

²⁷⁰ *Toshiba Am. Elec. Components, Inc. v. Lexar Media*, 124 Cal. App. 4th 762, 773 (App. Ct. 2004).

²⁷¹ *See id.* at 769 (“In some circumstances . . . principles of fundamental fairness require the demanding party to pay any significant ‘special attendant costs beyond those typically involved in responding to routine discovery.’”) (quoting *San Diego Unified Port Dist. v. Douglas E. Barnhart, Inc.*, 95 Cal. App. 4th 1400, 1405 (Cal. Ct. App. 2002)).

²⁷² *Lipco Elec. Corp. v. ASG Consulting Corp.*, No. 8775/01, 2004 WL 1949062, at *9 (N.Y. Sup. Ct. 2004). *See also* *Waltzer v. Tradescape & Co., L.L.C.*, 31 A.D.3d 302, 304 (N.Y. App. Div. 2006) (“[A]s a general rule, . . . the party seeking discovery should bear the cost incurred in the production of discovery material.”); *Rosado v. Mercedes-Benz of N. Am., Inc.*, 103 A.D.2d 395, 398 (N.Y. App. Div. 1984) (“each party should shoulder the initial burden of financing his own suit, and based upon such a principle, it is the party seeking discovery of documents who should pay the cost of their translation”).

discovery is willing to bear the cost of production of the electronic material.”²⁷³

[119] For example, in *Lipco Electrical Corp. v. ASG Consulting Corp.*, the plaintiff sought e-discovery that the defendant claimed would be “extremely difficult, time consuming and expensive” to extract from its computer hard drives or back-up tapes.²⁷⁴ Unable to find state court precedent dealing with discovery of electronic information, the court cited *Zubulake* for the proposition that electronic documents are discoverable.²⁷⁵ But the court then bypassed the federal analysis for deciding who should bear the cost of the discovery, finding that cost-shifting was not an issue because “under New York law, the party seeking discovery must bear the cost of production of the items for which discovery is sought.”²⁷⁶ Therefore, the court refused to order the production of the ESI at issue until such time as the requesting party expressed a willingness to pay for the associated costs, subject to later apportionment on proper application.²⁷⁷

[120] The *Lipco* court did not seem entirely comfortable with this outcome, noting that discovery of electronic information “raises a series of issues that were never envisioned by the drafters of the CPLR.”²⁷⁸ The court further noted that “[t]he cost of providing computer records can be rather substantial”²⁷⁹ and differs substantially from traditional paper discovery.²⁸⁰ The court’s analysis suggests that it might have reached a different conclusion but for the New York presumption in favor of discovery cost-shifting. Unless and until changes are made to the CPLR or New York’s highest court addresses this issue,²⁸¹ however, litigants in

²⁷³ *Lipco*, 2004 WL 1949062, at *9 (noting this was “especially true” where the requesting party “has been provided with hard copies of the electronically stored data”).

²⁷⁴ *Id.* at *6.

²⁷⁵ *See id.* at *7-8.

²⁷⁶ *Id.* at *9.

²⁷⁷ *See id.* at *10.

²⁷⁸ *Id.* at *6.

²⁷⁹ *Lipco Elec. Corp. v. ASG Consulting Corp.*, No. 8775/01, 2004 WL 1949062, at *8 (N.Y. Sup. Ct. 2004).

²⁸⁰ *See id.*

²⁸¹ The New York Administrative Board of the Courts has established statewide uniform rules, effective January 17, 2006, governing the jurisdiction and procedures of courts in

New York state court can expect that costs associated with obtaining discovery of electronic information will be borne by the party requesting the materials, and obtaining those materials may be conditioned on the requesting party affirmatively agreeing to pay such costs.²⁸²

(C) TEXAS AND MISSISSIPPI

[121] Texas and Mississippi have identical rules governing discovery of electronic information that mandate cost-shifting in certain circumstances.²⁸³ Those rules require that where the requesting party specifically asks for ESI, and specifies the form in which it should be produced, the responding party must produce what responsive ESI is reasonably available to it in the ordinary course of business.²⁸⁴ If the responding party cannot, through reasonable efforts, retrieve the data or

the Commercial Division of the Supreme Court, which handles complicated commercial cases. *See Administrative Order of the Chief Administrative Judge of the Courts*, Dec. 29, 2005, <http://www.nycourts.gov/comdiv/CD-Rules.pdf>. Among other things, the new rules address the topic of discovery of electronic information. Specifically, under Rule 8 of the new section 202.70 of the Uniform Rules for N.Y.S. Trial Courts, attorneys are required to meet and confer about discovery issues including data preservation plans, relevant data identification, data production, cost allocation, identification of individuals responsible for data preservation, confidentiality and privilege issues, and designation of experts. *See also* Uniform Rules for N.Y.S. Trial Courts, pt. 202, <http://www.courts.state.ny.us/rules/trialcourts/202.shtml#70>.

²⁸² A New York court may, under certain circumstances, require the producing party initially bear certain costs associated with discovery production. *See, e.g., Weiller v. N.Y. Life Ins. Co.*, No. 604285/04, 2004 WL 3245345, at *6 (N.Y. Sup. Ct. Mar. 16, 2005) (ordering party to initially bear costs associated with compliance with preservation order covering “all databases, electronic material, tape media, electronic media, hard drives, computer disks and documents” as related to certain categories of documents). While the *Weiller* court was “not insensitive to the cost entailed in electronic discovery” and suggested it would, “at the appropriate juncture, entertain an application by defendants to obligate plaintiff, the requesting party, to absorb all or part of the cost of the e-discovery it seeks,” it refused to “constrain the production of possibly relevant evidence on account of the later need to allocate the cost.” *Id.*

²⁸³ *See* TEX. R. CIV. P. 196.4; MISS. R. CIV. P. 26(b)(5). In Texas, the general rule is that “the expense of producing items will be borne by the responding party and the expense of inspecting, sampling, testing, photographing, and copying items produced will be borne by the requesting party” absent a court order “for good cause.” TEX. R. CIV. P. 196.6 (emphasis added).

²⁸⁴ *See* TEX. R. CIV. P. 196.4; MISS. R. CIV. P. 26(b)(5).

information requested, or produce it in the form requested, it can file a motion. If the motion is granted, the court must order the requesting party to pay for the reasonable expenses of any extraordinary steps required to retrieve and produce the information.²⁸⁵

[122] As the preceding discussion demonstrates, there is no “set in stone” standard or process for determining which ESI discovery expenses a court may be willing to shift to the requesting party and under what circumstances that shifting may occur. Accordingly, a practitioner contemplating a cost-shifting motion should be familiar not only with the standard from the relevant jurisdiction, but also other persuasive authority that may have broken new or analogous ground.

B. MOVING BEYOND BACK-UP TAPES: SHIFTING OTHER DISCOVERY COSTS

[123] While it is undisputed that the costs of restoring, translating, and electronically searching inaccessible data on back-up tapes are potentially eligible for shifting, those costs may represent a small portion of a client’s discovery budget.²⁸⁶ The million dollar question (literally) is when can the cost of attorney document review (which can skyrocket in discovery of ESI), and other expenses that identifying and producing ESI entail, be recovered from the requesting party?

²⁸⁵ See TEX. R. CIV. P. 196.4; MISS. R. CIV. P. 26(b)(5). See also *Bank of Am. Corp. v. SR Int’l Bus. Ins. Co.*, No. 05-CV5-5564, 2006 WL 3093174 (N.Y. Sup. Nov. 1, 2006) (“Texas Rule of Civil Procedure 196.4 . . . requires that whenever a court orders a responding party to produce information that is not ‘reasonably available,’ the court must require the requesting party to pay the ‘reasonable expenses of any extraordinary steps required to retrieve and produce the information.’”).

²⁸⁶ See *Zubulake III*, 216 F.R.D. 280, 289-90 (S.D.N.Y. 2003) (“[T]he estimated cost of restoring and searching the remaining backup tapes is \$165,954.67, while the estimated cost of producing them (restoration and searching costs plus attorney and paralegal costs) is \$ 273,649.39 (\$19,003.43 for the five sample tapes, or \$3,800.69 per tape, times seventy-two unrestored tapes), a difference of \$ 107,694.72.”).

1. THE COST OF REVIEWING DOCUMENTS FOR RELEVANCE, PRIVILEGE,
AND/OR OTHER PROTECTIONS

[124] The cost of reviewing and analyzing ESI for production can be a substantial item on a party's discovery bill. Privilege review is a difficult task in the paper world, and it quickly can become unmanageable when ESI is involved. One of the reasons ESI can hinder the discovery process and prove such a heavy draw on resources is that documents cannot just be handed over to the requesting party. Rather, documents must first be reviewed by attorneys for relevance and responsiveness and, even more importantly, to exclude, redact, and/or mark with appropriate confidentiality designations those documents that contain material that is subject to the attorney-client privilege and/or other protections (such as privacy and trade secrets). "The volume of such data, and the informality that attends the use of e-mail and some other types of electronically stored information, may make privilege determinations more difficult, and privilege review correspondingly more expensive and time-consuming."²⁸⁷ The task is further complicated by the existence of potentially privileged embedded data (e.g., earlier edits to a document) and metadata (e.g., identifying information automatically generated by word processing programs).²⁸⁸

[125] Despite the high costs of reviewing ESI for privilege, courts thus far have been reluctant to shift the cost of that burden to the requesting party, sometimes explaining that the producing party is uniquely positioned to control the scope of those costs.²⁸⁹ Contrary to that rationale, however, is

²⁸⁷ FED. R. CIV. PRO. 26(f) (amended 2006) comm. n.

²⁸⁸ *Id.*

²⁸⁹ See *Zubulake III*, 216 F.R.D. at 290 ("[T]he responding party should always bear the cost of reviewing and producing electronic data once it has been converted to an accessible form" because (1) "the producing party has the exclusive ability to control the cost of reviewing the documents" and (2) "the producing party unilaterally decides on the review protocol."); *Computer Assoc. Int'l, Inc. v. Quest Software, Inc.*, No. 02 C 4721, 2003 WL 21277129, at *2 (N.D. Ill. June 2, 2003) (ruling that costs of a third party computer consultant to make copies of eight imaged hard drives and search for privileged communications should not be shifted to the requesting party); *Byers v. Ill. State Police*, No. 99 C 8195, 2002 WL 1264004, at *12 (N.D. Ill. June 3, 2003) ("[P]laintiffs' motion is granted to the extent that they bear the cost of licensing the old e-mail program, though

the legal requirement that parties conduct a thorough review for privilege and similar protections or risk waiving the privilege/protection—not only for the actual documents produced, but possibly as to entire subject matters.²⁹⁰ Some courts have even held that a privilege review consisting solely of an electronic search for certain keywords is insufficient to preserve the privilege. Due to this conundrum, advocates should not unduly hesitate to argue that under Rules 26(b)(2)(C) and 26(c) the court should shift the costs of attorney review, or even deny particular discovery of certain electronic information altogether, in cases where the value of that discovery is outweighed by the burden of conducting an attorney review before production.²⁹¹ Logistically, courts could require the producing party to estimate the cost of reviewing ESI for privilege and work product just as they do for search and restoration, and could require a similar sampling procedure to provide a factual basis for the estimate.

2. OTHER ESI DISCOVERY COSTS TO CONSIDER SHIFTING

[126] The expansion of the scope of ESI discovery gives rise to and escalates a number of other expenses that should be considered for inclusion in cost-shifting requests, even though the courts have not yet addressed whether those costs shifting. Where the requesting party issues the typical, broad request for ESI and refuses to agree to accept anything

the defendants shall continue to bear the expense of any review for responsive documents, as well as for privileged or confidential material.”).

²⁹⁰ See, e.g., *Pogue v. Diabetes Treatment Ctrs. of Am.*, No. 99-3298, 2004 U.S. Dist. LEXIS 18763, at *6 (D.D. Cir. May 17, 2004) (stating that inadvertent disclosure waives the attorney-client privilege; “the scope of the waiver extends to all communications relating to the same subject matter, and will not be exempted, distinguished or balanced by ‘degrees of voluntariness,’ except in ‘extraordinary circumstances.’” (citing *In re Sealed Case*, 877 F.2d 976, 980 (D.C. Cir. 1981))).

²⁹¹ See, e.g., FED. R. CIV. PRO. 26(b)(2) (amended 2006) comm. note (“[T]he producing party’s burdens in reviewing the information for relevance and privilege may weigh against permitting the requested discovery.”); see, e.g., *In re General Instrument Corp. Sec. Litig.*, No. 96-C1129, 1999 WL 1072507, at *6 (N.D. Ill. Nov. 18, 1999) (non-monetary costs are properly considered in assessing burden imposed by discovery; “the technical matter of retrieving the documents from the backup tapes would be just the start of the process”). See also Laura E. Ellsworth & Robert Pass, *Cost Shifting in Electronic Discovery*, 5 SEDONA CONF. J. 125, 131 (2004) stating that (privilege review and production are as much costs of discovery as restoration and can be unduly burdensome or expensive because the volume of ESI is so much greater).

less than a burdensome, large-scale ESI preservation, collection, processing, review, and production effort, the responding party should consider whether to ask the court to shift the costs of those efforts to the extent the requested discovery is not denied outright.²⁹² Potential costs to be shifted include:

- Identifying all possible locations of discoverable data across a vast array of storage devices.
- Familiarizing counsel with IT systems to prepare for eventual collection and production.
- Preserving accessible data and halting a backup tape recycling program under a litigation hold.
- Productivity losses suffered when forensic operations make computer systems unusable.
- Disruption of normal business operations where billable in-house IT and other staff must attend to tasks related to ESI discovery.
- The cost of purchasing special processing and review tools to allow for the forensically sound migration and production of ESI.
- Processing large amounts of ESI to cull out data that does not meet basic relevance standards (e.g., date, file type, keywords, duplicates).
- Engaging experts, consultants, and specialists for tasks beyond restoration and translation (which are already considered shiftable), such as forensic collection of ESI, sampling, affidavits, and testimony.
- Producing ESI in the requested format.

[127] For example, the complexity of a client's network infrastructures and IT policies can make pinpointing the specific location of relevant ESI difficult and time consuming:

In today's increasingly digital workplaces, a small or medium-sized office may have five to 30 network servers that may be backed up to a single backup tape or backed up independently by server or function. Furthermore, larger

²⁹² See *infra* section IV.C.2.(d)

corporations typically operate by using hundreds or thousands of servers that may be segregated by function or task, business unit, or geography. Each of these servers may be backed up by a system of multiple backup tapes.²⁹³

Therefore, as discussed *supra* in Section III, one of the biggest challenges at the outset of a case is developing a thorough understanding of a client's network infrastructure and IT systems. Counsel must determine the client's capabilities to search for and retrieve electronic data, and understand what data may be accessible and what data is arguably inaccessible.

[128] Once located, the expense associated with preserving large amounts of accessible data can be enormous. *Zubulake* suggests that a potential litigant could preserve accessible electronic evidence by simply taking "a mirror-image of the computer system" as soon as the duty "to preserve documents in the state they existed at that time" attaches.²⁹⁴ "Taking a mirror-image of an entire, active computer system, however, in practice, likely will prove to be a more formidable task, particularly for large companies, companies with multiple or overseas offices, or companies that have employees that use various electronic data systems (laptops, home PC's, PDA's, etc.)."²⁹⁵ While *Zubulake* states that parties need only halt recycling of tapes that store documents of key players which are not otherwise available,²⁹⁶ the cost of acquiring new tapes for that category of data backup can mount quickly, as can the cost of locating the backup tapes that have the data of relevant custodians or for the relevant time frame.

[129] It also may be possible to recoup the cost of expert, consultant, and dedicated in-house staff time. For example, in *Portis v. City of Chicago*, the plaintiffs' law firm had compiled an electronic database of arrest

²⁹³ Sharp, *supra* note 269, at 34.

²⁹⁴ *Zubulake IV*, 220 F.R.D. 212, 218 (S.D.N.Y. 2003).

²⁹⁵ Stephen D. Whetstone & Kara A. Millonzi, *E-Discovery Document Preservation & Cost-Shifting: A Matter of Dollars and Sense* 2005, at 6, available at http://www.stratify.com/stratify_resources/articles/Whetstone_DocPreserv_and_CostShifting.pdf (last visited March. 19, 2007).

²⁹⁶ *Zubulake IV*, 220 F.R.D. at 218.

records from paper and electronic data produced by the City, a time-consuming and expensive project.²⁹⁷ After declining to participate in the compilation, the City decided it wanted to use the database, and it moved to compel production of the database. The court granted the motion, but ordered the City to split the plaintiffs' compilation costs, including the paralegals' and computer consultant's time spent on the project.²⁹⁸

[130] Finally, there is the cost of formatting ESI for production. Several cases that considered the production format for ESI before Amended Rule 34(b) gave the format choice to the demanding party. While a document request may still specify the form in which ESI is to be produced under Amended Rule 34(b), a responding party is also involved in determining the form of production because written responses must state the form the party intends to use for producing ESI if the requesting party does not specify a form or if the responding party objects to a form that the requesting party specified.^{FN} If the parties cannot agree on a production format and seek assistance in resolving the dispute, the court is not limited to the forms initially chosen by the requesting party, stated by the responding party, or specified in the rule.^{FN} Where the desired or ordered production format is burdensome to the producing party, shifting costs is one way to accommodate the burden.²⁹⁹

C. MAKING THE ARGUMENT TO SHIFT COSTS IN THE DISCOVERY OF ELECTRONIC INFORMATION

[131] The following discussion outlines some steps practitioners may find useful in developing and tailoring their cost-shifting arguments. While *Zubulake* is widely regarded as the most thorough consideration of cost-

²⁹⁷ *Portis v. City of Chicago*, No. 02 C 3139, 2004 WL 2512084, at *1 (N.D. Ill. Dec. 7, 2004).

²⁹⁸ *Id.* at *2-4. The plaintiffs calculated the cost based on the paralegals' and computer consultant's hourly billing rate, rather than the "out of pocket" rate actually paid by firm, i.e., the portion of their salaries actually paid by the firm for this work. The court held the billing rate was proper because it represented the plaintiffs' opportunity cost and put the parties in position they would be in had they collaborated on the project from the beginning. *Id.* at *4-6.

^{FN} FED. R. CIV. PRO. 34(b).

^{FN} FED. R. CIV. PRO. 34(b) (amended 2006) comm. n.

²⁹⁹ *See infra* at section IV(C)(2).

shifting to date, it is by no means settled law. Case law remains sparse, and attorneys may need to refer to other jurisdictions by analogy. Because the cost-shifting analysis is so fact-specific, courts develop and refine applicable standards with each new case on their dockets. The contours of successful cost-shifting arguments also vary by jurisdiction. This dynamic situation offers litigators a unique opportunity to advocate for cost allocations that are fair to their clients and helps further sound policy, while in the process shaping the evolution of discovery law.

1. GATHER INFORMATION EARLY: WHERE IS THE ESI, WHAT NEEDS TO BE DONE TO IT, AND WHAT WILL THAT COST?

[132] When faced with a burdensome discovery request with ESI discovery issues, one of the first tasks for counsel is to acquire an understanding of how much ESI exists and how much it will cost to produce. Regardless of the standard used, “the most important ingredient for the analytical process to produce a fair result is a particularization of the facts to support any challenge to discovery of electronic records.”³⁰⁰ Gathering the facts to support a cost-shifting argument requires investigating the full range of sources of ESI, any restoration and translation requirements, and which experts and specialists may be needed to complete the job and testify in court.

[133] The broad sweep of ESI that is potentially discoverable³⁰¹ may require a litigant to conduct a forensically sound collection of data in many formats from a multitude of storage devices. To gauge the potential cost of collection, counsel as well as appropriate business and IT personnel should survey the potentially relevant universe of data types and storage devices that parties may seek to discover, which could include:

³⁰⁰ *Thompson v. U. S. Dep’t. of Hous. & Urban Dev.*, 219 F.R.D. 93, 98 (D. Md. 2003).

³⁰¹ *See* FED. R. CIV. P. 34(a) (amended 2006) (stating that parties may inspect or otherwise access “writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations stored in any medium from which information can be obtained”). The Civil Rules Advisory Committee that drafted the e-discovery amendments noted that the amendment to Rule 34 was intended “to cover all current types of computer-based information” and “encompass future changes and developments.” FED. R. CIV. P. 34(a) (amended 2006) comm. n, cmt., n.

- Documents (including drafts), e-mails, instant messages, text messages, digitized voicemail, electronic calendars, web content (including intranets, extranets, blogs, web-based meetings, webcasts and podcasts), embedded and meta-data, fragmented and deleted data, software and source code.
- Desktop computers, laptop computers, network and e-mail servers, removable media (including compact discs, flash memory sticks, PDAs, iPods), cell phones, and archival and back-up media.

[134] The category of costs most likely to be shifted to the requesting party is that of searching and translating inaccessible electronic data into a usable form. The effort and expense involved in transforming ESI in various formats and locations into information searchable and readable by humans, can multiply quickly when that data is stored with no plan for retrieval in the future, let alone the targeted retrieval and review that litigation requires. Therefore, producing ESI is expensive, not only because of the additional time it requires, but also because of the specialized knowledge, services, and equipment of technical experts for everything from basic collection to testifying about the process in court.

[135] Additional ESI processing costs should be appraised, such as culling data down to a relevant time frame and relevant custodians, de-duping, key word searching, review by attorneys for relevance and privilege, and formatting ESI prior to production or for use as evidence at trial. Various document management and review tools may be required as each of these stages, which could involve licensing and other expenditures.

[136] Outside vendor, consultants, and expert expenses lie beyond the pale of ordinary business and litigation expenditures, so they make good candidates for cost-shifting. Counsel should assess categories of expert costs, including data collection, restoration and translation (including computer forensics), and expert testimony in court pertaining to discovery of electronic information issues. Additionally, keeping expert costs under control (e.g., insisting on a detailed budget, getting bids, monitoring costs) will enhance cost-shifting arguments to the court (and benefit attorney-client relationships where such arguments fail).

[137] ESI may be inaccessible or so broadly diffused across platforms and devices that it is difficult to locate. Therefore, the need for experts in the discovery of electronic information may arise as early as the data collection stage, and specialists may be required to identify, collect, and analyze run-of-the-mill records that another party requests. This is traditionally the work of lawyers or in-house personnel, and the addition of consultants and/or experts fundamentally changes the cost structure of litigation. Additionally, most litigants will need to retain outside vendors to perform the restoration and translation of inaccessible data. These tasks may include computer forensics work,³⁰² such as making mirror-images of hard drives and reconstructing lost, deleted, or damaged data.³⁰³ Restoration and translation predicates have no analog in traditional discovery, making them indisputable additions to the cost of discovery since the paper age. Furthermore, these are time-consuming, highly technical exercises that require many hours of work by expensive personnel.

[138] A litigant may also need to engage an expert witness to testify to the efficacy of its discovery technology, the projected cost of responding to a discovery request, the technical implications of various data retention programs, etc. Due to its technical nature, courts are increasingly calling for experts to testify in court regarding discovery requirements. An accusation of failure to preserve relevant ESI may lead the court to require

³⁰² Forensic examinations of hard drives and servers by outside experts raise additional, potentially costly issues: the risks of disclosing privileged material and of violating privacy laws. See GEORGE L. PAUL & BRUCE H. NEARON, *THE DISCOVERY REVOLUTION: E-DISCOVERY AMENDMENTS TO THE FEDERAL RULES OF CIVIL PROCEDURE* 142-44 (American Bar Association 2006). Because a forensics expert can access any data stored on the computer she investigates, counsel may need to devise ways to screen for privilege and privacy on the entire machine, regardless of potential responsiveness to the discovery request. This requirement imposes additional cost up front, as well as potential liability or loss of litigation leverage down the road as a result of waived privilege or disclosure of protected private data.

³⁰³ Courts may be more receptive to cost-shifting arguments where the other side requests discovery that can only be obtained by forensic work. For instance, in the employment case *Laurin v. Pokoik*, the plaintiff sought to discover the date certain ledger entries had been made in the defendant's computer, which could only be ascertained through forensics. In ruling on the request, the magistrate stated that the proper procedure would be for the plaintiff to move to compel at her own expense. *Laurin v. Pokoik*, No. 02Civ. 1938, 2004 WL 2724767, at *2 (S.D.N.Y. Nov. 30, 2004).

investigation by a neutral third-party expert. Again, these expenses exceed the traditional bounds of discovery. Attorneys should incorporate fees for testifying and investigatory experts into their discovery cost estimates early in the case, and advocate for cost-shifting wherever the benefit-burden analysis and fairness considerations require it.

[139] Finally, because discovery of electronic information raises a number of cost questions absent from traditional discovery, litigants hoping to resolve them in their favor should assess and address the above considerations as early in the case as possible. In this vein, the amended FRCP call for early discussion of discovery costs in several places, which can help lay the foundation for future cost-shifting arguments. Ultimately, a party that knows where its ESI is, what needs to be done to it, and how much that will cost, will be in the best position to urge a courts to shift the cost of some or all of its discovery costs.

2. LITIGATING THE BENEFITS AND BURDENS OF ESI DISCOVERY

[140] Under Rules 26(b)(2)(B) 26(b)(2)(C), and 26(c), the responding party may move for a protective order or the requesting party may move to compel when there is a dispute over the discovery of ESI. The FRCP and case law require courts resolving those disputes to balance the benefits of the ESI discovery at issue with the burden, including expense, on the producing party. The burden-benefit analysis is always a case-specific inquiry and a court's willingness to shift costs will vary with the issues at stake. For instance, drafts of contracts in a case may hold more benefit where intent is an issue than in a straight breach of contract case.³⁰⁴

[141] While discovery of electronic information can burden the producing party on a number of fronts, urging the court that *all* ESI discovery in a case is burdensome is not credible, as it is undisputed that computer technology can make the production of responsive information easier in some respects.³⁰⁵ Also, any *benefit* the discovery brings to the producing

³⁰⁴ See also *Cognex Corp. v. Electro Sci. Indus., Inc.*, No. Civ. A. 01CV10287RCL, 2002 WL 32309413, at *5 (D. Mass. July 2, 2002) (finding little benefit in recovering deleted e-mails in patent case as compared to an employment discrimination case).

³⁰⁵ See, e.g., *Zubulake I*, 217 F.R.D. 309, 318 (S.D.N.Y. 2003) (“[E]lectronic evidence is frequently cheaper and easier to produce than paper evidence because it can be searched

party is weighed against cost-shifting under all relevant tests: *Zubulake*, Rule 26(c), etc. Therefore, counsel for the responding party should assess any possible benefit to the requesting party in producing ESI and be prepared to modulate cost-shifting arguments accordingly in anticipation that the requesting party will point to those benefits to defuse requests to shoulder an ESI discovery bill.

[142] Shifting costs may be an appropriate balancing mechanism in a variety of circumstances, some of which are discussed below. In addition, effective cost-shifting arguments will highlight the legal and factual considerations that limit the likely benefits of the requested discovery. They also will distinguish unfavorable cases where the legal theories and/or factual circumstances at issue made the benefit to the requesting party greater, or the burden to the producing party lesser, than in the case at hand.

(A) THE “NOT REASONABLY ACCESSIBLE” THRESHOLD IS A MOVING TARGET

[143] Multiple information systems and constant technological change make the parameters of accessibility elusive.³⁰⁶ *Zubulake* attempted to break the world of ESI down into five categories (active, online data; near-line data; offline storage/archives; backup tapes; and erased, fragmented or damaged data) and then performed a cost-shifting analysis only as to the last two categories of “inaccessible” data³⁰⁷ Electronic data storage is more complicated and dynamic than this, however, a party need not

automatically, key words can be run for privilege checks, and the production can be made in electronic form, obviating the need for mass photocopying.”); *Itzenson v. Hartford Life and Accident Ins. Co.*, No. CIV. A. 99-4475, 2000 WL 1507422, at *2 (E.D. Pa. Oct 10, 2000) (rejecting defendant’s claim that searching thousands of electronic files for a particular subject would be unduly burdensome). *See generally* Shira A. Scheindlin & Jeffrey Rabkin, *Electronic Discovery in Federal Civil Litigation: Is Rule 34 Up to the Task?*, 41 B.C. L. REV. 327, 364 (2000) (“By comparison [to the time it would take to search through 100,000 pages of paper], the average office computer could search all of the documents for specific words or combinations of words in minute, perhaps less.”).

³⁰⁶ PAUL & NEARON, *supra* note 302, at 14 (“when the rules speak about whether something is ‘reasonable’ to do, the adjudication will depend more on the enveloping system than it will on the information itself”).

³⁰⁷ *Zubulake I*, 217 F.R.D. at 318-319.

abandon a cost-shifting argument simply because the discovery at issue does not fall into one of the two categories blessed by Zubulake.

[144] Courts generally agree that restoring archival or deleted data presents sufficient burden such as to warrant the consideration of cost-shifting.³⁰⁸ Litigants can generally term such ESI as “not reasonably accessible” without much controversy. Other categories are not so clear. The “not reasonably accessible” standard was *meant* to be a moving target. “It is not possible to define in a rule the different types of technological features that may affect the burdens and costs of accessing electronically stored information.”³⁰⁹ This flexibility can be used to try to convince the court that a cost-shifting argument succeeds on the particularized facts of a case.

(I) ACTIVE DATA MAY BE INACCESSIBLE FOR DISCOVERY PURPOSES.

[145] Even active ESI may pose undue burden and expense when being produced for litigation. The degree of burden and expense may depend on where the data resides. For example:

- Proprietary systems, used by many businesses from technology companies to accounting firms, may not have readily available searching and translation/production tools and requesting parties often cannot comprehend information produced from them in native format.^{FN}
- Relational databases, which exist in fluid form, must be manipulated to produce relevant, responsive information and can present potentially large confidentiality and privilege problems.³¹⁰

³⁰⁸ *E.g.*, *Antioch Co. v. Scrapbook Borders, Inc.*, 210 F.R.D. 645, 652 (D. Minn. 2002) (“[I]t is a well accepted proposition that deleted computer files, whether they be e-mails or otherwise, are discoverable”; ordering mirror imaging of defendant’s hard drive where plaintiff had agreed to pay cost.); *Zubulake III*, 216 F.R.D. 280, 291 (S.D.N.Y. 2003) (shifting some cost of restoring archival data).

³⁰⁹ FED. R. CIV. P. 26(b)(2) (amended 2006) comm. n.

^{FN} Amended Rule 34(a) does require that, if necessary, a responding party “translate” information it produces into a “reasonably usable” form. The Committee Note to Rule 34(b) further states that “[u]nder some circumstances, the responding party may need to

- Computer systems that are significantly more complicated than typical systems found a most businesses—for example where the company recently has undergone numerous mergers or acquisitions and thus data on certain areas of the system are extremely difficult to access—may be held to be inaccessible due to the system’s extraordinary “complexity.”³¹¹

[146] Each of these actively used data sources from which data would be difficult or expensive to extract could present strong arguments for cost-shifting under Rule 26(b)(2)(B) based on burden or expense.

(II) BUSINESSES SHOULD BE FREE TO CHANGE INFORMATION SYSTEMS
WITHOUT BEING PENALIZED IN LATER LITIGATION.

[147] The FRCP Amendments recognize that technology will continue to evolve, and leave the benefit-burden standard flexible to accommodate that change.³¹² Nonetheless, pre-amendment case law suggests that courts may decline to deem ESI “not reasonably accessible” where the inaccessibility results from the client’s choice to decommission a system. For example, in *Xpedior Credit Trust v. Credit Suisse First Boston (USA)*,

provide some reasonable amount of technical support, information on application software, or other reasonable assistance to enable the requesting party to use the information.” FED. R. CIV. P. 34(b) (amended 2006) comm. n.

³¹⁰ *E.g.*, *Multitech. Services, L.P. v. Verizon S.W.*, No. CIV A 4:02-CV-702-Y, 2004 WL 1553480 (N.D. Tex. July 12, 2004) (holding that while a report requested from defendant’s client databases was accessible, the plaintiff has to pay half the cost of creating it. Finding *Zubulake* inapposite, the court nonetheless ran through the factors and concluded that splitting the expense provided the defendant with incentive to control costs while also recognizing that only the plaintiff would benefit from the report’s creation and production).

³¹¹ PAUL AND NEARON, *supra* note 301, at 130 (2006) (“[W]hat the committee is addressing [in Rule 26(b)(2)(B)] is . . . the difficulty of access because of the complexity of a system.”) (emphasis omitted).

^{FN} *See id.* at 84-87 (summarizing numerous “negative” comments submitted in response to a discussion draft of amended Rule 34 regarding production in “native format,” including a lengthy comment submitted by Microsoft).

³¹² *See* FED. R. CIV. P. 26(b)(2) (amended 2006) comm. N., cmt. n. (“It is not possible to define in a rule the different types of technological features that may affect the burdens and costs of accessing electronically stored information.”).

Inc., Judge Scheindlin (of *Zubulake* fame) applied her cost-shifting factors to archival optical discs that had been made inaccessible when the defendant merged with another company and decommissioned the Unix servers that accessed the discs.³¹³ With four factors weighing against cost-shifting and the rest neutral, the court denied costs even though the defendant no longer had the hardware and software needed to access data.³¹⁴

[148] There are strong arguments that *Xpedior* is has been superseded by Amended Rule 26(b)(2), which states that a responding party does not have to produce ESI from sources that are “not reasonably accessible because of undue burden or cost” While potential litigants may have *some* duty to maintain reasonable access to ESI expected to be relevant to litigation within a reasonable future time period, the force behind the Amendments and current court decisions is to find fair ways to allocate discovery costs in a rapidly changing technology environment. Inevitably and by their nature, electronic storage platforms and media will come and go, a fact that should be considered by courts in judging accessibility.

(III) ALTERNATIVE SOURCES CAN PRECLUDE THE NEED TO PRODUCE
INACCESSIBLE ESI.

[149] Responding parties can challenge requests to produce inaccessible ESI and avoid the resulting cost-shifting spats if the information is available elsewhere in other formats. “In many circumstances the requesting party should obtain and evaluate the information from [reasonably accessible] sources before insisting that the responding party search and produce information contained on sources that are not reasonably accessible.”³¹⁵

³¹³ *Xpedior Creditor Trust v. Credit Suisse First Boston (USA), Inc.*, 309 F. Supp. 2d 459, 465 (S. D.N.Y. 2003).

³¹⁴ *Id.* at 467.

³¹⁵ FED. R. CIV. P. 26(b)(2) (amended 2006) comm. n., cmt. n.

[150] In seeking to show that information requested from inaccessible sources is available elsewhere, counsel should consider questions such as:

- Is there evidence that relevant data has been not been deleted, and thus is available from accessible sources?
- Is there a company policy requiring hard copy print-outs of the inaccessible data?³¹⁶
- Can the client identify paper duplicates of the inaccessible electronic files?
- Is the information available from an outside source, such as the client's auditors?
-

If information is available elsewhere at less expense, the requesting party should have to pay the cost of acquiring it for more.³¹⁷

(IV) CAUTION: REGULATORY RETENTION REQUIREMENTS MAY SERVE AS ACCESSIBILITY BENCHMARKS.

[151] Clients subject to document retention requirements by regulatory agencies such as the IRS or SEC may find that courts presume the accessibility of any ESI falling within those requirements.³¹⁸ This issue is best addressed with clients *before* litigation. A robust data lifecycle management policy³¹⁹ will not only comply with those regulations but also ensure the most-efficient and easiest access possible to electronic data that is required to be retained. Most companies face some sort of retention mandate, and are best advised not to have the only source of such ESI be a pile of unmarked backup tapes, as a court is unlikely to entertain a cost-shifting argument in this instance. On the other hand, regulatory requirements may create alternative sources, such as auditors or examiners, of ESI that the party has otherwise destroyed. A court may

³¹⁶ See, e.g., *Rowe Entm't, Inc. v. Williams Morris Agency, Inc.*, 205 F.R.D. 421, 428 (S.D.N.Y. 2002) (attempting to show alternative sources failed where no company policy required hard copies). Nothing in Amended Rule 34(b), prevents the court from ordering that paper is a "reasonably usable" form where the only other source for the information is "not reasonably accessible."

³¹⁷ *Id.* at 430.

³¹⁸ See PAUL & NEARON, *supra* note 301, at 139-41.

³¹⁹ See *supra* section III.E.

require a party to go to those sources for electronic data no longer internally accessible to a responding party before it will consider whether data must be restored under Rule 26(b)(2) and at whose cost.

(B) SAMPLING ITSELF CAN PRESENT AN UNDUE BURDEN

[152] The Committee Note to Amended Rule 26(b)(2)(B) contemplates sampling in certain circumstances and courts will likely see a trend of increasing requests to order use of that methodology for cost-shifting analysis or otherwise. While a potentially powerful cost-cutting tool, sampling also can raise its own issues of burden and cost depending on the circumstances. When a responding party asserts that electronic data is not reasonably accessible, “[t]he requesting party may need discovery to test this assertion,” which “might take the form of requiring the responding party to conduct a sampling of information contained” on the inaccessible sources.³²⁰ Nearly all courts considering accessibility and whether to shift costs have required sampling of the data in question,³²¹ and amended Rule 34 explicitly permits the inspection, testing, and sampling of ESI.³²² The purpose of sampling is two-fold: to estimate the cost of producing the full request, and to develop the factual record sufficiently to narrow that request.

[153] Sampling alone may be unduly burdensome for the responding party: “where the cost of a sample restoration is significant compared to the value of the suit, or where the suit itself is patently frivolous, even this minor effort may be inappropriate.”³²³ Parties may raise the burden or intrusiveness of sampling under Amended Rules 26(b)(2) and 26(c).³²⁴ Conversely, sampling can serve the responding party’s interests where it provides an opportunity to educate the court on the breathtaking expense and limited value of what the other party has asked for. Hard data

³²⁰ FED. R. CIV. P. 26(b)(2), (amended 2006) advisory comm. note, cmt. n.

³²¹ *E.g.*, *Zubulake I*, 217 F.R.D. 309, 324 (S.D.N.Y. 2003); *Rowe*, 205 F.R.D. at 432.

³²² *See* FED. R. CIV. P. 34(a) (amended 2006) comm. n., cmt. n. Rule 34 was changed “to make clear that parties may request an opportunity to test or sample materials sought under the rule.” *Id.*

³²³ *Zubulake I*, 217 F.R.D. at 324 n.77.

³²⁴ FED. R. CIV. P. 34(a) (amended 2006) comm. n., cmt. n.

showing the court that the requesting party seeks an “awfully expensive needle to justify searching a haystack” could do just that.³²⁵

[154] There are of course other ways to develop the factual record sufficiently to allow a court to rule on cost-shifting questions. The parties are limited only “by the court’s own imagination and the quality and quantity of factual information provided by the parties to be used by the court in evaluating the Rule 26(b)(2) factors.”³²⁶ Armed with the detailed understanding provided by the fact-gathering process and through discovery or meet-and-confer conferences, counsel can encourage the court to rule on the facts before it without the expense and delay of sampling. Advocates also can forestall sampling with issue-narrowing motions.³²⁷

(C) THE AMOUNT OF DISCOVERY EXPENSE THAT WILL TRIGGER COST-SHIFTING IS FACT-SPECIFIC

[155] The case-specific nature of the benefit-burden analysis suggests that almost any ESI discovery expense may make a good candidate for cost-shifting under Rule 26(b)(2), depending on the circumstances of the case. As a result, the amount of out-of-pocket discovery expense that will convince a court to shift costs may vary widely. The Supreme Court has generally directed courts to look at “whether the cost is substantial[,] not whether it is ‘modest’ in relation to ability to pay.”³²⁸ For example, in the discovery of electronic information context, a New York district court shifted the mere \$1,680 it would cost one of the defendants to create a special program to extract documents from inaccessible sources.³²⁹

³²⁵ See *McPeck v. Ashcroft*, 202 F.R.D. 31, 34 (D.D.C. 2001).

³²⁶ *Thompson v. U.S. Dept. of Hous. & Urban Dev.*, 219 F.R.D. 93, 99 (D. Md. 2003).

³²⁷ *E.g.*, *Chudasama v. Mazda Motor Corp.*, 123 F.3d 1353, 1365 (11th Cir. 1997) (holding the court has power to stay discovery where it may be mooted by pending, potentially dispositive motions on issues to which discovery is directed).

³²⁸ *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 361-62 (1978) (holding that a \$16,000 expense to compile class member list was “substantial” and should be shifted to plaintiff, who sought the information).

³²⁹ *Anti-Monopoly, Inc. v. Hasbro*, No. 94 Civ. 2120, 1996 WL 22976, at *2 (S.D.N.Y. Jan. 23, 1996).

[156] When the court evaluates burden by comparing the cost of production to the amount in controversy, that will make costs less likely to shift in a high-stakes case (e.g., patent litigation) than in a smaller-stakes case (e.g., a contract dispute) unless the costs at issue are high. For example, the court in *Zubulake* concluded that a factor of “several fold” between the cost of production and the stakes (\$273,650 to produce versus a potential multi-million dollar recovery) weighed against cost-shifting, noting that in a multi-million dollar case, “the cost of restoration is surely not ‘significantly disproportionate’ to the projected value” of the case.³³⁰ Litigants in high-stakes cases where the amount of alleged damages is speculative, will be subject to significant dispute, and will depend on expert testimony should argue against the use of “amount in controversy” as the basis for evaluating whether burden is undue.

(D) COST-SHIFTING IS APPROPRIATE WHERE DISCOVERY REQUESTS ARE TOO BROAD

[157] As ESI has raised the specter of endlessly broad discovery, courts have used cost-shifting to help parties think twice about the information they seek. “Where a party multiplies the litigation costs by seeking expansive rather than targeted discovery, that party should bear the expense.”³³¹ This standard can be particularly useful for parties that receive egregiously broad discovery requests that are typical in certain types of complex litigation. By the same token, the more specific a discovery request is, the more reasonable it remains for the responding party to pay its own expenses.³³² In this vein, the end game of a cost-shifting argument may be not a reimbursement check, but a narrower discovery request.

[158] Courts likely will be friendly to arguments that couch cost-shifting in the context of encouraging thoughtful, targeted discovery. Further,

³³⁰ *Zubulake III*, 216 F.R.D. 280, 288 (S.D.N.Y. 2003).

³³¹ *Rowe Entm’t, Inc., v. William Morris Agency*, 205 F.R.D. 421, 430 (S.D.N.Y. 2002).

³³² *See Zubulake III*, 216 F.R.D. at 284 (requesting that all e-mails from five UBS employees during two-year period was sufficiently specific to weigh against full cost-shifting, so court shifting only 25% of cost to requesting party); *see also Zubulake I*, 217 F.R.D. 309, 321 (S.D.N.Y. 2003) (“Specificity is surely the touchstone of any good discovery request.”).

clients may prefer to pay to respond to a narrowly tailored request than have the hollow victory of reimbursement for an exhaustive production that disrupts normal business. The threat of cost-shifting as a pressure to narrow discovery requests also has good consequences at both the practical and policy levels. Practically speaking, narrower discovery promotes judicial efficiency and reduces costs to both parties. From a policy perspective, requiring a requesting party to pay for broad discovery helps to prevent fishing expeditions and the use of discovery as a sword to induce settlement.

[159] When collecting ESI, there are various options to narrow the scope of discovery requests, some of which are generally applicable to the discovery context and some of which are specific to discovery of electronic information. They include:

- Urging the requesting party to reword or drop vague qualifiers, such as those calling for “all documents relating to” a particular subject matter;
- Limiting the custodians, dates, sources, and locations of electronic data to be searched;
- Using keyword searches and other culling tools; and
- Providing an initial “sample” of documents as the starting point for negotiations.³³³

[160] Advocates who are fluent in the methods and technologies available to narrow discovery requests and reduce the expense of production will be more effective in resolving cost issues both for meet-and-confer efforts and for the court. Counsel and their clients should also agree on how narrowly the requests should be drawn to warrant a compromise on costs.

(E) PRODUCTION FORMATS MAY RAISE ADDITIONAL BURDEN ISSUES

[161] The format for production is an important feature of the production of ESI. Requesting parties now may specify the desired production format in discovery requests to “facilitate the orderly, efficient, and cost-effective

³³³ See *supra*, section III.

discovery of electronically stored information.”³³⁴ Though the requesting party may now specify a preferred format, the responding party should argue against producing formats that incur unreasonable costs. If a dispute over format is brought to the court for resolution, the responding party should consider arguing that to the extent the court orders production in a burdensome format the requesting party should shoulder some or all of the extra burden involved in producing in that format.

[162] Before meeting and conferring on discovery, counsel should become familiar with the technical and cost implications to the particular client of producing in the various electronic formats that the other side may request. The categories of electronic formats, from least to most potentially problematic to produce, include:

- Images (most commonly .tiff or .pdf files; may be electronic or paper printouts);
- Native format (requires appropriate software on requester’s part), with or without metadata;
- Copies of backup or archival media (with software to access, or in readable form);
- Mirror images of hard drives; and
- Direct access to computer systems.

[163] Where the requesting party does not specify a form, Rule 34(b) allows the responding party to produce ESI in reasonably useful form *or* in the form in which it is “ordinarily maintained.”³³⁵ The Advisory Committee note to Rule 34(b) clarifies that the Rule requires the responding party to “translate” data that is not “reasonably usable.”³³⁶ Under some circumstances, the responding party “may need to provide some reasonable amount of technical support, information on application software, or other reasonable assistance to enable the requesting party to use the information.”³³⁷ Responding parties should consider arguing

³³⁴ FED. R. CIV. P. 34(b) (amended 2006) comm. n.

³³⁵ FED. R. CIV. P. 34(b) (amended 2006).

³³⁶ FED. R. CIV. P. 34(b) (amended 2006) comm. n.

³³⁷ *Id.*

against these burdensome forms of production under the benefit-burden test of Rule 26 and asking for cost-shifting in the event a burdensome form of production is required.³³⁸

[164] Native format production has been endorsed by a number of courts.³³⁹ Producing in native forma, however, requires additional technical efforts—such as locking data so the requesting party cannot alter it and a more complex method of electronic Bates stamping—and additional privilege and relevance review of metadata, deleted data, and fragmented data, which may in turn require further technical measures to effect electronic redaction. There are also well-documented problems with the use of native format files as evidence in litigation. Responding parties should resist a format that would require production in more than one form as contrary to Rule 34(b)(iii), but it is difficult to conceive of how “native” files can be used in depositions, to support motions, or at trial without further processing and duplicative production.

3. BALANCED VERSUS LOPSIDED CASES

[165] Since cost-shifting involves a case-specific analysis, advocates must tailor arguments on the issue to suit the case at hand. In particular, litigants may face different discovery dynamics depending on whether a case is “balanced” (parties with roughly equal resources and discovery

³³⁸ *Id.* (“The questions whether a producing party should be required to convert [not reasonably usable data] to a more usable form, or should be required to produce it at all, should be addressed under Rule 26(b)(2)(B)”). For examples of recent district court decisions approving lower-cost production formats, *see* *Zakre v. Norddeutsche Landesbank*, No. 03 Civ. 0257, 2004 WL 764895, at *1 (S.D.N.Y. Apr. 9, 2004) (approving production of 204,000 emails in text-readable format without any further relevance cut where this was “in as close a form as possible as they are kept in the usual course of business”); *In re Lorazepan and Clorazepate Antitrust Litig.*, 300 F. Supp. 2d 43, 47 (D.D.C. 2004) (stating that no index was required if CD-ROMs were searchable; “[t]he 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”).

³³⁹ *E.g.*, *Hagenbuch v. 3B6 Technologies, LLC*, No. 04 C 3109, 2006 WL 665005 (N.D. Ill. Mar. 8, 2006); *Williams v. Sprint/United Mgmt Co.*, 230 F.R.D. 640 (D. Kan. 2005); *In re Verisign, Inc. Securities Litig.*, No. C 02-02270 JW, 2004 WL 2445243 (N.D. Cal. Mar. 10, 2004); *In re Telxon Corp. Sec. Litig.*, No. 5:98CV2876, 2004 WL 3192729 (N.D. Ohio July 16, 2004).

needs), or “lopsided” (parties with significantly different resources and discovery needs). These differing dynamics may require different discovery strategies and cost-shifting arguments.

[166] Courts will consider the parties’ resources under any of the cost-shifting tests. In a balanced case, the relative equality in resources weighs for cost-shifting in either direction.³⁴⁰ Therefore, any expansive discovery request made by a party in a relatively balanced case will (1) be made by the other side in return, (2) put the party at risk for defending against a cost-shifting request, and (3) make the court less sympathetic to the party’s own cost-shifting arguments for the other side’s requests. Resources of course weigh against cost-shifting when the requesting party is significantly smaller than the party requesting to have costs shifted. This is no reason, though, not to argue for cost-shifting, cost-sharing, or denial of discovery. In *Zubulake*, for instance, the court found that the fact that the plaintiff was an individual and UBS a large corporation weighed against cost-shifting. The court nonetheless went on to consider the plaintiff’s personal wealth and her lawyers’ ability to advance costs, and assigned 25% of the costs of electronic data recovery to her.³⁴¹

[167] Where the requesting party appears to be disadvantaged by size, advocates for the responding party should remind the court that size disparity alone does not justify burdensome discovery and that the fairest solution may be to deny the request or require that it be narrowed significantly. Where discovery is allowed, the respective parties’ resources should not be the controlling factor in the cost-shifting determination. It costs almost nothing for a requesting party to put together and issue an extremely broad, highly burdensome request for production of documents. Even a “small” requesting party can gain a significant strategic advantage by sending out a boilerplate request, which then effectively puts the onus on the responding party to either spend time and money preserving, collecting, processing, reviewing, and producing enormous volumes of ESI early on in the case or spend time and money

³⁴⁰ See, e.g., *OpenTV v. Liberate Techs.*, 219 F.R.D. 474, 478 (D. Cal. 2003) (taking into account parties’ similar in situation in patent litigation when splitting the cost of source code extraction).

³⁴¹ *Zubulake III*, 216 F.R.D. 280, 288 (S.D.N.Y. 2003).

litigating discovery disputes. This “Hobson’s Choice” allows the requesting party to abuse the discovery process by putting this pressure on the responding party, even where it is “big,” to settle the matter rather than litigate on the merits.

[168] This dynamic is exacerbated by the fact that in “lopsided” cases a “small” requesting party typically has little or no ESI to be produced in the litigation. That means that the requesting party has no incentive to be reasonable in fashioning and negotiating the scope of requests for production of ESI—it will have no need for the responding party to be reasonable in kind as to its own requests.

[169] In many types of litigation, the cost of ESI production in “lopsided” cases has caused the playing field to be imbalanced throughout the discovery process, in favor of the “small” requesting party with no ESI of its own to produce. Only when courts take all of the relevant factors into consideration in “lopsided” cases can the shifting of certain costs of ESI production to the requesting party truly begin to balance that playing field. It may be necessary for courts to begin taking a very different approach to the shifting and sharing of ESI discovery costs to bring modern-day, complex litigation back into alignment with the mandate of FRCP 1.

4. REQUESTING THAT PARTICULAR DISCOVERY BE DENIED

[170] It is important to remember that cost-shifting arguments are only a back-up to asking the court to deny burdensome discovery requests outright. In some cases, outright denial may be more appropriate protection from burdensome discovery requests than a cost-shifting order. Attorneys should not become so enamored with the idea of shifting discovery costs that they neglect to consider protesting the underlying request in the first instance. Rule 26(b)(2)(C) and 26(c) provide that a court may modify *or deny* a discovery request if the requesting party has had ample opportunity to obtain discovery, if the material sought is available elsewhere, or if the request is unduly burdensome. Two cases show district courts deciding to deny discovery instead of shifting costs, based on those policy grounds.

[171] In *In re General Instrument Corporate Securities Litigation*, the court denied discovery of e-mails from back-up tapes even though production would not be unduly expensive, because the projected benefit failed to outweigh the burden and ample opportunity for discovery had been availed.³⁴² The defendant had already produced 110,000 pages from one year of back-up tapes when the plaintiffs moved to compel further production, despite having stated at an earlier status conference that discovery had concluded. As to the benefit, the court found that the large number of documents already produced made it unlikely that additional documents were necessary, and that the plaintiffs had failed to identify “any specific factual issue” that would make it so.³⁴³ As to the burden, the court found that it would be significant given the volume of the e-mail at issue and the necessity of reviewing it for privilege and responsiveness.³⁴⁴ With expert discovery beginning, the court concluded that “[f]orcing defense counsel to engage in document review would necessarily distract their energies from the other parts of this ongoing litigation.”³⁴⁵

[172] In *Cognex Corp. v. Electro Scientific Industries, Inc.*, the court denied plaintiffs’ motion to compel restoration and production of data from 820 back-up tapes covering a period of nine years, even where plaintiffs offered to pay the full cost and even though such a broad search would inevitably unearth relevant documents.³⁴⁶ For one, the defendant had “already conducted an extensive search for relevant documents” and produced thirty boxes of documents from every employee who had worked on the product in dispute.³⁴⁷ “At some point, the adversary system needs to say ‘enough is enough’ and recognize that the cost of seeking every relevant piece of discovery is not reasonable.”³⁴⁸ The court also found there would be limited benefit derived from the discovery, distinguishing the patent infringement case before it from employment

³⁴² *In re General Instrument Sec. Litig.*, No. 96 C 1129, 1999 WL 1072507, at *5-6 (N.D. Ill. Nov. 18, 1999).

³⁴³ *Id.* at *6.

³⁴⁴ *Id.*

³⁴⁵ *Id.*

³⁴⁶ *Cognex Corp. v. Electro Sci. Indus., Inc.*, No. Civ. A. 01CV10287RCL, 2002 WL 32309413, at *5 (D. Mass. July 2, 2002).

³⁴⁷ *Id.* at *1.

³⁴⁸ *Id.* at *5.

discrimination cases where relevant e-mails might be deleted by guilty individuals but remain on backup tapes.³⁴⁹ Finally, noting that willingness to pay does not waive limits on the number of interrogatories and depositions a party can take, the court concluded that “[t]here is something inconsistent with our notions of fairness to allow one party to obtain a heightened level of discovery because it is willing to pay for it,” and that “[t]he sense of fairness underpinning our system of justice will not be enhanced by the courts participating in giving strategic advantage to those with deeper pockets.”³⁵⁰

[173] The outright denial of superfluous or unduly burdensome discovery requests serves FRCP 1’s policy goals of speedy cost-efficient resolution. It also prevents the potential unfairness of allowing parties to “buy” disproportionate discovery, which would represent a serious mishandling of the cost-shifting principles inherent in Rule 26 and applied in federal case law.³⁵¹

V. CONCLUSION

[174] Advances in litigation technologies provide a great deal of promise for litigants. Those technologies are still in their relative infancies, and we are only beginning to scratch the surface of their potential to limit the costs of discovery. As concept and other forms of searching methodologies mature and become more cost-effective, it is likely that they will become more streamlined and accepted by courts. In light of the growing complexity of the world of ESI, innovation and aggressive advocacy for use of those technologies is needed to improve the efficiency of the discovery process, and not just at the review and production stages. The increasing need to make early decisions about the scope of relevant documents during preservation efforts warrants the use of efficient technologies at that stage as well.

³⁴⁹ *Id.*

³⁵⁰ *Id.*

³⁵¹ FED. R. CIV. P. 1; *see also* FED. R. CIV. P. 26(b)(2)(b) comm. n. (“A requesting party’s willingness to share or bear the access costs may be weighed by the court in determining whether there is good cause [for production].”)

[175] While the courts are slowly moving away from the presumption that the producing party pays for the cost of production, litigants need to become better versed at articulating in plain English the unique burdens involved in locating and producing ESI. Cost-shifting rules remain jurisdiction specific, with some forums more receptive to cost-shifting arguments than others. However, the arsenal of arguments to be considered is wide-ranging and should provide some measure of relief for clients facing enormous discovery bills.

[176] Finally, litigants should be aggressive in invoking FRCP 1 as a basis for the innovative use of search strategies and cost-shifting to increase efficiency and reduce costs across the board in discovery. It is only in this way that the mandate of a just, speedy, and inexpensive determination of every action will become a reality in discovery.