

Civil Rules Discovery Subcommittee Mini-Conference at DFW Airport
September 9, 2011

Preservation, Search Technology & Rulemaking

Thomas Y. Allman, Jason R. Baron and Maura R. Grossman¹

I. Introduction

The Rules Committee has sought information about and input on the influence of technology – including predictable future developments – on the possible rulemaking needed to govern preservation obligations. As broadly defined, various forms of automated technologies in addition to search technology are implicated by the question.²

The purpose of this relatively brief Essay is to highlight certain “hot button” issues arising with respect to automated versus manual search methods and technologies, specifically with respect to their current and future use in meeting the initial duty to preserve electronically stored information (ESI). Included in our discussion are certain “cutting edge” techniques that are advocated as effective in identifying preservable information in diverse storage applications throughout the enterprise.³

By way of background, we first describe the role of search technology generally, before turning to the preservation context and our evaluation of the need for rulemaking on the topic. In our view, the drafters of the 2006 Federal

¹ Mr. Allman, an Attorney and former General Counsel, is a Chair Emeritus of The Sedona Conference® Working Group on Electronic Document Retention and Production (WG1). Mr. Baron, Director of Litigation at the National Archives and Records Administration, is one of the current Co-Chairs of WG1. Maura R. Grossman, Counsel at Wachtell, Lipton, Rosen & Katz, is a member of WG1. The views expressed herein are those of the authors alone and have not been reviewed, endorsed or authorized by any public or private institution (or client) with which they are affiliated, including The Sedona Conference®. The authors wish to thank the in-house and outside counsel who reviewed earlier drafts and contributed significantly to our articulation of the issues, as well as Kevin Cheng, a legal intern in Mr. Baron’s office, for his research assistance.

² The Sedona Conference® Survey of WG1 Members (August 2011) (copy on file with authors) makes reference to centralized enterprise search, email journaling, and indexing and collection software, among others.

³ Adam Cohen, *Angst Over Data Retention Ruling May be Misplaced*, 8/24/2009 Nat’l L. J. 15 (Col. 1) (noting products which build on centralized management of data, including enterprise search tools and compliance monitoring tools), copy at <http://www.law.com/jsp/law/LawArticleFriendly.jsp?id=1202433204035>.

Amendments wisely did their best to promulgate “technology neutral” approaches to solving e-discovery issues, and the same result should obtain in 2011.

Review for Responsiveness

The status quo ante consists of information being identified and preserved in response to potential or pending litigation followed, when necessary,⁴ by collection, culling, processing and review for relevance and privilege. The latter steps in the process – uniformly regarded as the most costly of the e-discovery workflow due to the involvement of counsel in the process – increasingly have been subject to search-technology enhancement. Whether such methods can be said to be successfully utilized at the earlier stages of preservation and collection remains a more open question.

Despite its limitations,⁵ key word searching, using simple words or word combinations, with or without Boolean operators, is “[b]y far the most commonly used” methodology in the filtering of data for production of responsive information in discovery.⁶ However, alternative search techniques, taking advantage of “predictive coding,” concept searching, and other forms of machine learning, are increasingly used to prioritize and select documents for review.⁷ These techniques are backed up by quality control measures, sampling, and informed project-oriented management.⁸

Recent studies suggest that appropriate use of these techniques can yield results that are superior to exhaustive manual review,⁹ as measured by “recall” and

⁴ There is anecdotal evidence that the majority of litigation holds do not result in collection of the data placed on hold – the holds are primarily prophylactic. However, the “collection” process is sometimes incorporated into and made part of the initial identification and preservation process itself.

⁵ The Sedona Conference® Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 8 SEDONA CONF. J. 189, 201 (2007) (*Sedona Search Commentary*) (“simple keyword searches end up being both over- and under- inclusive in light of the inherent malleability and ambiguity of spoken and written English (as well as all other languages)”).

⁶ *Id.* at 200.

⁷ *Id.* at 203 (“Anecdotal information suggests that a small number of companies and law firms . . . are using alternative search methods to either identify responsive documents (reducing expensive attorney review time) or to winnow collections to the key documents for depositions, pretrial pleadings, and trial”); *see generally*, Jason R. Baron, Law in the Age of Exabytes: Some Further Thoughts on ‘Information Inflation’ and Current Issues in E-Discovery Search, 17 RICH. J.L. & TECH 9, at *30 et seq. (2011) (discussing “predictive coding”).

⁸ The Sedona Conference® Commentary on Achieving Quality in the E-Discovery Process, 10 SEDONA CONF. J. 299 (2009)(encouraging greater use of project management, sampling and other means of verifying the accuracy of the e-discovery process).

⁹ *See, e.g.*, Maura R. Grossman and Gordon V. Cormack, Technology-Assisted review in E-Discovery Can Be More Effective and Efficient Than Exhaustive Manual Review, 17 RICH. J.L. & TECH 11, at *1-*2 (Spring 2011) (results of one study found that “a technology-assisted process, in which humans examine only a small fraction of

“precision,” i.e., how effective a given method is in finding all relevant documents, and how accurate it is in eliminating “false positive,” or nonrelevant materials, respectively.

As the *Sedona Search Commentary* states in Practice Point 1, “[i]n many settings involving electronically stored information, reliance *solely* on a manual search process for the purpose of finding responsive documents may be infeasible or unwarranted. In such cases, the use of automated search methods should be viewed as reasonable, valuable, and even necessary.”¹⁰

Nevertheless, even with the most advanced automated techniques, it has become clear that some level of manual review – at initial stages of coding, as a quality control check throughout, and especially for privilege – remains an important part of the workflow process designed to assure that relevant and non-privileged material is identified and produced. We also readily acknowledge that in smaller cases, traditional manual review may continue to constitute the primary means for accomplishing the review task.

The Federal Rules and the accompanying Committee Notes do not address or mandate any particular review methodology nor limit the use of technology in its implementation. Courts have correctly concluded that there is no obligation to “examine every scrap of paper in its potentially voluminous files,” and have cited Sedona Principle 11¹¹ in support of the use of “reasonable selection criteria,” such as search terms or samples to access and identify “potentially responsive electronic data and documents.”¹²

More recently, in connection with privilege review issues, the Evidence Advisory Committee has noted that advanced search techniques may play a role in

the document collection, can yield higher recall and/or precisions than an exhaustive manual review process, in which humans code and examine the entire document collection”); Patrick Oot et al., *Mandating Reasonableness in a Reasonable Inquiry*, 87 DENV. U. L. REV. 533, 551 (2010) (“use of auto-categorization systems can potentially reduce document request response times from over four months to as little as thirty days for even the largest datasets.”); Bennett Borden, *E-Discovery Alert: The Demise of Linear Review* (2010)(alternative search techniques can facilitate faster review times), copy at http://www.clearwellsystems.com/e-discovery-blog/wp-content/uploads/2010/12/E-Discovery_10-05-2010_Linear-Review_1.pdf.

¹⁰ *Sedona Search Commentary*, 8 SEDONA CONF. J. at 208 (italics added).

¹¹ Principle 11, *The Sedona Conference Best Practices Recommendations & Principles for Addressing Electronic Document Production* (“*The Sedona Principles*”)(2nd Ed. June 2007)(A responding party may satisfy its good faith obligation to preserve and produce relevant [ESI] by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information.).

¹² *Treppel v. Biovail*, 233 F.R.D. 363, 374 (S.D.N.Y. 2006).

the context of avoiding a finding of privilege waiver.¹³ The *Victor Stanley I* opinion strongly advocated the application of such advanced techniques to future reviews for responsiveness and privilege.¹⁴

Identification for Preservation

In contrast, the identification of information subject to preservation often must be planned and executed without the benefit of precise knowledge of potential discovery issues. The duty to preserve may arise even before litigation is filed, or before counsel for the requesting party is identifiable – and certainly before the Rule 26(b) conference. It is not surprising, therefore, that the FJC Survey presented at the Duke Conference showed limited use of the conference for that purpose. Thus, initial preservation decisions are often made unilaterally,¹⁵ and a party must take into account the uncertainty as to eventual discovery.¹⁶ Thus, preservation may involve retention of broad categories of sources (such as key and ancillary custodians), or searches of potential sources for subject matter information within a given time frame or on a specific topic.

Automated search techniques may be used for targeted or selective identification from sources such as archives or LAN servers. Increasingly, it is also argued by vendors that the ability to “index” the contents of diverse information sources permits centralized search for and identification of information responsive to legal holds in multiple sources.¹⁷

¹³ FRE 502 Explanatory Note (Revised 11/28/2007)(“Depending upon the circumstances, a party that uses advanced analytical software applications and linguistic tools in screening for privilege and work product may be found to have taken ‘reasonable steps’ to prevent inadvertent disclosure. The implementation of an efficient system of records management before litigation may also be relevant.”).

¹⁴ *Victor Stanley v. Creative Pipe*, 250 F.R.D. 251, 259 n.9 (D. Md. 2008)(noting existence of “other search and information retrieval methodologies [which] include: probabilistic search models, including ‘Bayesian classifiers,’” as well as “Fuzzy Search Models,” “Clustering” searches and “Concept and Categorization Tool[s].”).

¹⁵ Thomas Y. Allman, *Managing Preservation Obligations After the 2006 Federal E-Discovery Amendments*, 13 RICH. J.L. & TECH. 9, at *26 (2007) (“absent agreement with opposing counsel, unilateral preservation decisions about inaccessible sources always carry some risk of post-production challenge for potential spoliation”).

¹⁶ *See, e.g.*, *Wm. T. Thompson v. General Nutrition*, 593 F. Supp. 1443, 1455 (C.D. Cal. 1984) (the duty to preserve attaches to what 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”).

¹⁷ Data Sheet, Symantec Enterprise Vault™ Discovery Collector (2010)(“[e]xtend[ing] indexing and classification of data beyond the archive to include a full spectrum of enterprise sources – network servers, storage systems, application repositories, and personal computers across the enterprise as well as remote locations”), copy at http://eval.symantec.com/mktginfo/enterprise/fact_sheets/b-ev_9_discovery_collector_DS_20982272-2.en-us.pdf.

These techniques are said to enable a party to “crawl”¹⁸ across diverse data sources in order to identify content in repositories subject to hold criteria, regardless of custodian or source.¹⁹ Once identified, the material can be locked down in place via a “hold procedure,” or transferred electronically to secure storage pending review and production.²⁰ The concept of “reaching in” to a variety of indexed content silos,²¹ or to material in the “cloud,”²² bears a resemblance to an earlier suggestion by one court that a party might meet preservation obligations by “conducting system-wide keyword searching and preserving a copy of each ‘hit.’”²³

Advocates for this approach argue that such an enterprise-wide search can achieve better results than the “unpredictability and inconsistency of self-collection.”²⁴

II. Preservation Today

In meeting preservation responsibilities, a party need extend only reasonable and good faith efforts, proportionate to the issues and risks involved, as not “every conceivable step” is required.²⁵ The Sedona Commentary on Proportionality²⁶ explains that the “burdens and costs of preservation” of potentially relevant information should be “weighed” when determining the “appropriate scope of

¹⁸ Charles Babcock, Oracle Challenges Google With New Enterprise Search Engine, Information Journal (March 2006)(describing how search engine can “crawl through file systems and the databases that underlie” various applications and databases as well as email systems and documents, web servers or other sources).

¹⁹ Data Sheet, Autonomy Control Point (2011)(“information is indexed automatically making it visible, transparent, and available to be controlled and governed”), copy at

<http://protect.autonomy.com/products/compliance/controlpoint/index.htm> (scroll to “related products”).

²⁰ Barry Murphy, In-Place Preservation – A Workable Solution?, eDiscovery Journal (2010)(corporations aspire to “in-place preservation” under which “files are placed on legal hold and locked down where they live instead of being copied and moved to a specific preservation repository” which, inter alia, “reduces overall risk because there is less data movement”), copy at <http://ediscoveryjournal.com/2010/11/in-place-preservation-a-workable-solution/>.

²¹ See Andrew Cohen, EMC White Paper (April 2006)(“Federated Search” is the capability of automating search and retrieval of content from various sources within an enterprise by “reaching into” these various applications), copy at <http://www.emc.com/collateral/software/white-papers/h2153-prac-ent-meth-compl-ediscovery-wp.pdf>.

²² Data Sheet, Clearwell Identification and Collection Module (2011)(describing “federated search-enabled collection” of material in “Microsoft Exchange, Microsoft SharePoint and Windows files shares whether deployed on-premise or in the cloud”), copy at http://www.clearwellsystems.com/ediscovery-news/pr_06_13_11.php.

²³ *Sedona Search Commentary, supra*, 8 SEDONA CONF. J. 189 at 200 (citing to, *inter alia*, Zubulake v. UBS Warburg LLC, 229 F.R.D. 422, 432 (S.D.N.Y. 2004)).

²⁴ Data Sheet, Recommind Accelerate ESCA & Collection (2011)(enables enterprises to “identify, preserve [and] collect”), copy at http://www.recommind.com/products/accelerate_discovery/eca_and_collection.

²⁵ Principle 5, *The Sedona Principles* (The obligation to preserve [ESI] requires reasonable and good faith efforts to retain information that may be relevant to pending or threatened litigation. However, it is unreasonable to expect parties to take every conceivable step to preserve all potentially relevant [ESI].).

²⁶ The Sedona Conference® Commentary on Proportionality, 11 SEDONA CONF. J. 289 (2010).

preservation.” Thus, transient or ephemeral data that is not kept in the ordinary course of business and that the organization may have no means to preserve need not be preserved under normal circumstances.²⁷

Traditionally, the decision on what documents and data to preserve has been left to the informed judgments of custodians, assisted, as appropriate, by counsel, and the IT department. This approach is said to be used by “[a] majority of organizations.”²⁸

The Traditional Approach

The focus in pre-discovery preservation of ESI is on user-created or “unstructured” information residing in email, electronic documents, spreadsheets and other similar materials, as well as structured data in the form of databases. It is preservation of the unstructured data, however, which presents the most challenges – and leads to the most disputes in the reported sanction decisions.²⁹

Unstructured information is typically found in active files stored on servers, laptops or office desktops, or other distributed sources (including removable media). It may also be found in third-party cloud-based storage which is susceptible to the control of the entity. It may take the form of email and attachments, compressed and encrypted email archives, spreadsheets, text messages,³⁰ tweets, instant message (IM) chats, or information available on social networks.

The preservation process typically begins with issuance of a litigation hold, triggered by the onset or anticipation of litigation. As described in *Zubulake IV* “[o]nce a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy and put in place a ‘litigation hold’ to ensure

²⁷ *Columbia Pictures v. Bunnell*, 2007 WL 208419 at *3-6 (C.D. May 29, 2007), review denied, 245 F.R.D. 443 (2007).

²⁸ According to William Tolson, “[a] majority of organizations still rely on the practice of instructing custodians to search for and protect potentially responsive ESI locally.” Post, *Are Custodial Self-Discovery and Preserving ESI in Place Good for You?* (January 19, 2011), copy at <http://blog.ironmountain.com/2011/compliance/are-custodial-self-discovery-and-preserving-esi-in-place-good-for-you/>.

²⁹ Dan H. Willoughby, Jr. et al., *Sanctions for E-Discovery Violations: By The Numbers*, 60 DUKE L. J. 789, 803 (2010) (“[i]n the 230 cases in which sanctions were awarded [in the study], the most common misconduct was failure to preserve ESI”).

³⁰ *See, e.g.*, Computer Forensics and E-discovery, Text message Usage is Exploding, January 2010, copy at <http://www.fulcrum.com/text-messages.htm> (noting that text message system typically bypasses a business entity’s central server/storage and backup processes and “text message content typically exists only in the senders’ and recipients’ devices”).

the preservation of relevant documents.” Use of a litigation hold was acknowledged in the Committee Notes to the 2006 Amendments,³¹ and its implementation is covered by the recently amended Sedona Commentary on Legal Holds.³²

A litigation hold notice is typically directed to pertinent custodians to retain potentially relevant documents, including ESI and, in some cases, seeking certification that they have taken steps to ensure that it has not been destroyed.³³ The form of the hold may vary according to the circumstances.³⁴ It typically spells out the reasons for the hold and lists the topics subject to it, as well as the manner in which identified information is to be handled. It may ask targeted custodians to identify other potential custodians of potentially relevant data. There may or may not be automated processes in place to track issuance of the litigation hold and to record communications regarding compliance.

The custodian is often responsible for identifying and preserving information stored on the “endpoint devices” he or she uses, such as desktops, laptops and removable devices. Depending upon the specificity of the litigation hold, there could be some selectivity involved in applying the criteria. In many (but not all) cases, the information is then collected for purposes of responding to discovery requests, often without any specific attempt to winnow or cull the information prior to institution of the review process.

The IT department and, in some cases, counsel, may play a role, depending on the scope of the preservation effort. IT is usually responsible for accessing enterprise systems such as databases and implementing any affirmative actions required to support preservation activities. Selective backup media might or might not be retained, depending upon the likelihood that it captured unique copies of relevant materials.³⁵ LAN drive information as well as hard drives from desktops or laptops of former employees who were potentially involved might be retained if not already redeployed. Procedures to address computer maintenance and repair activities for custodians on holds often are also considered.

³¹ Comm. Note, Rule 37(f)(2006)(“intervention in the routine operation of an information system is one aspect of what is often called a ‘litigation hold’”).

³² The Sedona Conference® Commentary on Legal Holds: The Trigger & The Process (2nd Ed. 2011), 11 SEDONA CONF. J. 265, 286-287 (2010).

³³ Mayer Brown Slides, A Refined Litigation Hold (2010), 10-12, copy at http://www.mayerbrown.com/public_docs/12-01-10_Comm_Lit_Webinar_E-Discovery_Slides.pdf.

³⁴ There are circumstances where oral notice may be sufficient. *See* n. 40, *infra*.

³⁵ *See, e.g.,* Wells Fargo Bank v. LaSalle Bank, 2009 WL 2243854, at *3 (S.D. Ohio July 24, 2009) (refusing to order restoration of backup media where hard copies of important emails were retained in loan files).

If incoming and outgoing email has been routinely archived through message journaling, it may or may not be decided to “execute a hold search” at that time to identify email within the archive subject to the hold.³⁶ In some cases, multiple keyword searches may be necessary to fully execute litigation holds against other data storage silos.³⁷

As recognized by some courts³⁸ and commentators, there are potential limitations on custodian-centric approaches to meeting a party’s preservation duty.³⁹ These include the problem of inconsistent, idiosyncratic methods for preserving ESI; late identification of key evidence; the possibility of metadata spoliation; the issue of self-interest or bias on the part of the end-user charged with the task; the non-lawyers absence of legal knowledge, including as to relevancy; and a general failure of attorneys to adequately supervise the process where it involves multiple (and sometimes huge numbers of) would-be custodians.

However, the issue is highly fact-specific, and in some contexts it can be quite reasonable to rely upon the assistance of custodians in selecting material

³⁶ This may depend upon the likelihood that the contents of the archive are subject to culling or application of retention periods by policy in the absence of such a hold. *See, e.g.,* Velocity Press v Bank, 2011 WL 1584720, at *1 (D. Utah April 26, 2011)(emails and attachments archived for one year unless a litigation hold is applied).

³⁷ Mia Mazza, et al., In Pursuit of FRCP 1: Creative Approaches to Cutting and Shifting the Costs of Discovery of Electronically Stored Information (hereinafter “Creative Approaches”), 13 RICH. J.L. & TECH. 11, at *34 & *79 (noting possible uses in what is “fast becoming a very complex and costly stage of discovery”).

³⁸ For example, in Pension Committee of the University of Montreal Pension Plan v. Banc of America, the Court found plaintiffs’ litigation hold policy defective in part because:

“It does not direct employees to preserve all relevant records-both paper and electronic-nor does it create a mechanism for collecting the preserved records so that they can be searched by someone other than the employee. Rather, the directive places total reliance on the employee to search and select what that employee believed to be responsive records without any supervision from Counsel.”

685 F.Supp.2d 456, 473 (S.D.N.Y. 2010) (as amended May 28, 2010); *see* Adams v. Dell, 621 F. Supp.2d 1173, 1194 (D. Utah 2009) (holding that defendant had violated its duty to preserve information, in part because the defendant’s preservation practices “place operations-level employees in the position of deciding what information is relevant”); Jones v. Bremen High School, 2010 WL 216640, at *7 (N.D. Ill. May 25, 2010)(non-lawyers do not have enough knowledge to correctly recognize which documents are relevant and otherwise may fail to reveal their own mistakes or misdeeds).

³⁹ *See, e.g.,* ARMA, “Is ‘Manual’ Collection of ESI Defensible? (4/10), copy at <http://www.arma.org/news/enewsletters/index.cfm?ID=4270>; Dean Gonsowski, Clearwell Systems, “Adams v. Dell Questions Custodian-Based Retention” (5/28/09), copy at <http://www.clearwellsystems.com/e-discovery-blog/2009/05/28/adams-v-dell-questions-custodian-based-retention-and-litigation-hold-practices-in-electronic-discovery/>; ; John Wang, “Automated Collection: Mitigating the Risks and Costs of Manual Collection” (8/12/10), copy at <http://grokify.com/2010/08/12/automated-collection-mitigating-the-risks-and-costs-of-manual-collection/>; James Shook, “Weekend at Bernie’s and End-user based eDiscovery” (8/10), copy at <http://www.kazeon.com/blog/2010/08/weekend-at-bernie-s-and-end-user-based-ediscovery/>

subject to a litigation hold,⁴⁰ given their greater familiarity with the specific language used and the methods and locations of retention. In addition, if it is not deemed to be feasible to achieve satisfactory results, other methods are available to supplement custodian-based preservation.⁴¹

For example, copies might also be made of specific custodians' mailboxes and files from active drives and other networked shared sites. Backup tape rotations may be modified so as to retain potentially relevant backups. In addition, a forensic image can be made of the desktop environment to remove the element of risk that deleted information could escape preservation.

One key issue, regardless of the form of identification, is whether to leave the information in place (i.e., on live networks), or to undertake its collection and storage, for potential use in future discovery. Preservation in place has, however, been subject to criticism.⁴²

Collectively, these concerns point towards counsel being more actively involved in ensuring that thoroughness in preservation and collection is achieved. However, as one of the authors has pointed out elsewhere,⁴³ the specific role of retained counsel in implementing a team-based approach is determined by the party, upon whom the obligation to preserve lies.⁴⁴ In any event, a party should

⁴⁰ In cases where broad categories of information are sought to be placed on hold or where relative small numbers of key custodians are involved, reliance on custodial collection – even on oral instructions – can be reasonable. *See, e.g., Orbit One v. Numerex*, 271 F.R.D. 429, 441 (S.D.N.Y. 2010) (“[i]n a small enterprise, issuing a written litigation hold may not only be unnecessary, but it could be counterproductive, since such a hold would likely be more general and less tailored to individual records custodians than oral directives could be”).

⁴¹ Greg Buckles, Don't Give Up on Custodial Self Collection, *edisccovery Journal* (5/25/2011) (“[w]e should not throw out the baby with the bathwater, just because [a Court] has come forth with an example of poorly executed discovery”), copy at <http://edisccoveryjournal.com/2010/06/don%E2%80%99t-give-up-on-custodial-self-collection/>.

⁴² *See, e.g., Brandon D'Agostino, Is the Use of “Preserve in Place” a Gamble in Electronic Discovery* (2010) (pointing out the volatility of storage media, the disruption to business continuity where ESI is locked down, and discussing expense and overpreservation), copy at <http://www.clearwellsystems.com/e-discovery-blog/2010/11/30/is-the-use-of-%E2%80%9Cpreserve-in-place%E2%80%9D-a-gamble-in-electronic-discovery/>; Albert Barsocchini, *Preserve in Place vs. Collect to Preserve*, *Inside Counsel* (Sept. 2009) (pointing out the danger of relying on users who may have proxy rights to delete data, the possibility of metadata failing to be preserved, and the general problem of incompleteness if an entity has continuing preservation duties that carry forward into the future), copy at <http://www.insidecounsel.com/2009/08/24/preserve-in-place-vs-collect-to-preserve>.

⁴³ Thomas Y. Allman, *Deterring E-Discovery Misconduct With Counsel Sanctions: The Unintended Consequences of Qualcomm v. Broadcom*, 118 *Yale L.J. Pocket Part* 161, 164 (2009) (“A client is ethically entitled to limit the responsibility of retained counsel in regard to a discovery engagement, which may well occur when teams of internal experts and vendors are involved”).

⁴⁴ *Compare Casale v. Kelly*, 710 F. Supp. 2d. 347, 365 (S.D.N.Y. April 26, 2010) (“responsibility for adherence to the duty to preserve lies not only with the parties but also, to a significant extent, with their counsel”) *with Centrifugal Force v. Softnet Comm.*, 2011 WL 1792047, at *3 (S.D.N.Y. May 11, 2011) (the obligation to preserve evidence is placed by the Second Circuit “on the ‘party,’ not on counsel” and is met if the party has taken reasonable steps).

work with its IT staff in fashioning ways to work within existing platforms and networks to more efficiently preserve and collect ESI across the enterprise.

A cautionary note about the use of technology in preservation is in order, however, as described next.

Future Developments

First, we believe that there are dangers lurking in over-reliance on “state of the art” automated technologies, such as “predictive coding,” in attempting to completely satisfy a party’s early preservation obligations. The proven efficacy of predictive coding for purposes of early case assessment and document review notwithstanding, such techniques simply remain unproven at this time in addressing the more comprehensive obligation to save ESI for preservation,⁴⁵ and thus may raise defensibility red flags if and when challenged.

Second, the capabilities of automated technology to enable search of the indexed content of multiple storage silos is subject to extravagant and largely unproven claims. The purported advantages include an enhanced ability to manage the repositories pursuant to policy and to avoid the “save-everything” mentality.⁴⁶ Some of the offerings also assert a capability to “automatically update the hold” as the data is revised or new data is added.⁴⁷

There is little publicly available information about the enterprise search approach, although one commentator describes it as a “pro-active” approach which is “now a reality, and is used by an increasing number of firms to prepare for litigation.”⁴⁸ There are, however, knowledgeable skeptics based on the costs and practicability issues involved.⁴⁹

⁴⁵ See Farrah Pepper, “Robot Review: Will Predictive Coding Win the Trust of Courts?,” Law Technology News (Aug. 1, 2011), copy at

<http://www.law.com/jsp/lawtechnologynews/PubArticleLTN.jsp?id=1202508331112&slreturn=1&hbxlogin=1>.

⁴⁶ Andrew Cohen, *supra*, EMC White Paper at 19 (advocating use of a “matter vault” for resulting materials within the content management environment for the life of the case).

⁴⁷ Exterro Partner Brief (2010), copy at <http://www.exterro.com/wp-content/uploads/2011/02/Recommind-and-Exterro-Partner-Brief.pdf>.

⁴⁸ Michael D. Berman, et al, Has Indexing Technology Made Zubulake Less Relevant?, ABA Section of Litigation Newsletter (Feb. 11, 2010), copy at <http://www.esi-mediation.com/pdf/hasIndexingMadeZubulakeLessRelevant.pdf>.

⁴⁹ Joe Dysart, Discovery In-House Approach, ABA Journal, August 2011, 32 (“The cost to a large organization, or probably even a smaller or medium organization, to index everything – well, not only is it untenable, but it’s probably cost-prohibitive” [quoting the Director of Legal and Compliance at JPMorgan Chase & Co.]), copy at http://www.abajournal.com/magazine/article/in-house_approach_corporate_lawyers_say_e-discovery_software_savings_there/.

For example, it has been suggested that “the reality of poor connectivity, slow storage, highly mobile decision makers and the radical growth of corporate ESI have kept this promise [enterprise-wide indexing and search] from becoming reality for most corporations.”⁵⁰ Other serious impediments include the very real limits raised by concerns involving inter-connection or control of related corporate entities. It is also possible that significant barriers may be created by the existence of multi-national data storage in countries subject to strict data privacy barriers.⁵¹

Whatever one’s level of optimism that in the near future, successful methods of enterprise search will emerge, it remains the case that at least for today, counsel’s active involvement in fashioning clear and consistent guidance for custodians to implement, coupled with greater involvement of IT staff, constitutes best practice in this area.

III. Rulemaking

The 2006 Amendments suggest that parties to civil actions in Federal Courts should discuss preservation at the Rule 26(f) conference, in order to arrive at practical agreements on scope, timing, and the mechanics to be employed in carrying out a legal hold.

The authors strongly endorse this approach,⁵² when feasible, given the need for a change in culture to deal with the profound and irrevocable changes involving information.⁵³ The Sedona Conference® Cooperation Proclamation is supportive of this effort as well. In many cases, parties can successfully agree on the practical limitations on custodians, date ranges and search and retrieval methodology for the preservation stage.

However, we do not recommend that Rule 26 or its Committee Note be amended to indicate any preference for a particular methodology or technology to be used in implementing preservation obligations. It is simply not possible to

⁵⁰ Greg Buckles, Desktop Collection 2.0 – Tackling the Enterprise Part 1, eDiscovery Journal, (2010), copy at <http://ediscoveryjournal.com/2010/06/desktop-collection-2-0-%E2%80%93-tackling-the-enterprise-part-1/> (“[t]hey have the illusion of live enterprise search, but only as long as they do not look too closely at search results”).

⁵¹ One experienced E-Discovery In-House Manager, who has been pitched on the topic, notes that this includes overcoming barriers to accomplish searches on widely distributed information systems which are not centrally accessible as well as limits on the ability of these tools to access encrypted files (such as Lotus Notes archives) or other files with password protections.

⁵² See, e.g., Allman, Managing Preservation Obligations, *supra*, 13 RICH. J.L. & TECH. 9, at *18 (“[b]oth parties should discuss preservation steps already undertaken and any plans for intervention in business processes”).

⁵³ George L. Paul and Jason R. Baron, Information Inflation: Can the Legal System Adapt?, 13 RICH. J.L. & TECH. 10, at *3 (“Litigators must collaborate far more than they have in the past, particularly concerning the discovery of information systems”).

anticipate the rapidly changing needs to which clients and their counsel they may need to respond.

As one Respondent in the poll of The Sedona Conference® WG1 Members put it, “[t]he growing dominance of the cloud, handhelds and social networking were almost entirely out-of-mind in 2006. Why would any thinking person assume that we are done with development of new and innovative ways to create, communicate and store information?”⁵⁴

The Rules Committee has correctly refused to “take sides,” or to require expenditures or investments in technology – the rules deal only with outcomes and leave to parties the determination of the reasonable steps needed to comply.⁵⁵ Thus, while there may be advantages to centralization of search techniques for purposes of preservation, it should not become an obligation under the Federal Rules any more than it already is under the common law.⁵⁶ Sedona Principle 6 correctly posits that the choice of “procedures, methodologies, and technologies appropriate for preserving and producing” ESI should be, absent agreement, made by the party in the best position to assess its own capabilities.⁵⁷

The goal, after all, is not perfection in executing preservation, but rather, making a reasonable effort.⁵⁸ Smaller organizations, for example, “may meet the requirements” of their committed compliance efforts with “less formality and fewer resources than would be expected of large organizations.”⁵⁹

⁵⁴ The Sedona Conference® Survey of WG1 Members (August 2011)(copy on file with authors)(a rule that employs specific directives will be “as out-of-touch as the 2006 FRCP amendments would have been if the drafters had lacked the wisdom to shy away from citation of specific technologies”).

⁵⁵ *See, e.g.,* *Paramount Pictures v. Replay TV*, 2002 WL 32151632, at *3 (C.D. Cal. May 30, 2002)(refusing to order development of software so as to facilitate collection of ephemeral information); *accord* *Oppenheimer Fund v. Sanders*, 437 U.S. 340 (1978)(it “borders on the frivolous” to argue that a party must keep records in a manner most convenient for future litigants).

⁵⁶ *Cache La Poudre Feeds v. Land O’Lakes*, 244 F.R.D. 614, 628 (D. Colo. 2007) (refusing to mandate an enterprise wide keyword search of multiple information sources).

⁵⁷ Principle 6, *The Sedona Principles* (“Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information”).

⁵⁸ *Valentin v. Bank of New York Mellon*, 2011 WL 1466122, at *2 (S.D.N.Y. April 14, 2011) (search methodology needed where there is a large volumes of data is not necessarily appropriate in cases with smaller volumes since reasonable steps – not perfection – is the goal).

⁵⁹ Commentary to the Sentencing Guidelines for the United States Courts, Title 18, Effective Compliance and Ethics Program, 18 USCS Appx. §8B2.1(C)(iii)(“In appropriate circumstances, reliance on existing resources and simple systems can demonstrate a degree of commitment that, for a large organization, would only be demonstrate through more formally planned and implemented systems.”).

The Federal Rules should, however, continue to encourage parties to develop reasonable policies and practices to manage ESI throughout its life cycle.⁶⁰ An innovation of the 2006 Amendments – Rule 37(e) – was intended to provide reassurance to those executing preservation obligations in good faith that a loss through routine processes would not be sanctionable. One of the authors has advocated that this provision be strengthened and broadened.⁶¹

At the very least, existing provisions could be strengthened through acknowledgement in advisory notes and in more detailed local rules that encourage parties to consider the use of innovative techniques across the entire e-discovery spectrum, including at the preservation stage.

III. Conclusion

As one of the authors recently noted, the “re-engineering [of] the discovery process [is] playing out against the backdrop of profound, transformational change.”⁶² The authors have no doubt that a bright future exists for the legal profession in harnessing the power of artificial intelligence and other 21st century information retrieval methods, so as to accomplish the goals of Federal Rule 1. However, it would be premature to assume that alternative search methods, including but not limited to “predictive coding,” are mature enough to be used in a defensible manner for purposes of meeting preservation obligations.

Moreover, while increased centralization of data management by entities may create opportunities to assist in preserving “low-hanging fruit” through enterprise search, the prospect of its availability does not justify abandonment of settled preservation practices, including appropriate custodian-based collection, as described above.

⁶⁰ Arthur Andersen v. United States, 544 U.S. 696, 704 (2005) (acknowledging destruction as a normal aspect of management of information); Sedona Conference® Commentary on Email Management: Guidelines for the Selection of Retention Policy, 8 SEDONA CONF. J. 239 (2007) (describing alternative approaches).

⁶¹ Connecticut recently amended its Rules to provide this approach. See Sec. 13-14 Connecticut Practice Book (2011)(eff. Jan. 2012)(limiting sanctions “for failure to provide information, including electronically stored information, lost as the result of the routine, good-faith operation of a system or process in the absence of a showing of intentional actions designed to avoid known preservation obligations”), copy at http://www.jud.ct.gov/Publications/PracticeBook/PB_070511.pdf.

⁶² J. Baron, Law in the Age of Exabytes, *supra*, 17 RICH. J.L. & TECH. 9, at *3.