

# **Search Method in E-Discovery: How Rule 26's Silence Poses a Risk of Sanctions to Attorneys and Increases the Cost of Litigation**

**KHANH T. HUYNH**

7 U. MASS. L. REV. 236

## **ABSTRACT**

The 2006 Amendments to the Federal Rules of Civil Procedure are the first codified references in the FRCP to electronic discovery. However, the lack of comprehensive rules in this area provides opportunities for attorneys to leverage search terms as a weapon, primarily to wear out opponents financially. Disagreement on search terms used to produce documents can prolong litigation. Complicated Boolean search terms can be difficult to run. Other search methods, such as natural language search, cannot provide efficient and accurate results. The cost to run complicated searches is high, and the lack of rules addressing search terms in the FRCP leaves parties at risk of sanctions. In addition, since electronic information may be extracted from third parties' servers, disagreement on search terms may lead to opponents using data extracted from third parties' as evidence of a failure to produce, and seeking sanctions. The low efficacy of search methods currently used also poses a risk of under-producing to attorneys. This article proposes two amendments to FRCP 26. First, FRCP 26 should mandate an agreement on search methodology and storage spaces to be searched. Second, FRCP 26 should allow electronic discovery production certification as complete and accurate using the agreed upon search method in agreed upon storage spaces.

## **AUTHOR**

Candidate for Juris Doctor, California Western School of Law, 2012. Many thanks to the UMass Law Review editorial team for their great work on my note. I would like to thank Professor Kenneth Klein for his guidance and Professor Kathryn Fehrman for teaching me legal writing. Thanks to Susan Sindelar for her excellent edits and encouragement, to my mother Muoi Chau and my partner Lori Glatzel for their love and support. Any remaining mistakes are mine alone.

**CONTENTS**

<b>I. INTRODUCTION.....</b>	<b>238</b>
<b>II. HOW A RULE AMENDMENT CAN HELP .....</b>	<b>241</b>
<b>III. THE SCOPE AND COST OF SEARCHING ELECTRONICALLY STORED INFORMATION .....</b>	<b>244</b>
<b>IV. TECHNICAL DIFFICULTIES AND OTHER PROBLEMS IN USING KEYWORD SEARCH TERMS TO FIND RESPONSIVE DOCUMENTS ...</b>	<b>249</b>
<b>V. KEYWORD SEARCHING IS REALLY A PROBLEM AND NOT A SOLUTION .....</b>	<b>250</b>
<b>VI. THE COURT’S AWKWARD POSITION IN THE SEARCH TERMS WAR.....</b>	<b>254</b>
<b>VII. THE CALL FOR COOPERATION .....</b>	<b>255</b>
<b>VIII. OTHER SOLUTIONS ATTEMPTED BY THE COURT AND ATTORNEYS .....</b>	<b>257</b>
<b>A. Technical Master.....</b>	<b>257</b>
<b>B. Other Search Methods.....</b>	<b>259</b>
1. Natural Language Search.....	260
2. Concept Search .....	261
<b>IX. CONCLUSION .....</b>	<b>262</b>

## I. INTRODUCTION

*“They kept asking me – did you attend this conference in 1993? That was fifteen years ago! How can I know which conference I went to fifteen years ago?”*

*- Jennifer Isaac, Project Manager,  
Qualcomm, Inc.<sup>1</sup>*

At that moment, Ms. Isaac did not know the answer,<sup>2</sup> however, the deposing attorney who was questioning her knew. Using the right keywords, that attorney had unearthed emails showing Qualcomm employees’ attendance at that 1993 conference. This information was later used to show an alleged scheme to conceal information, leading to enormous sanctions and potential disbarment. The intersection of contemporary electronically stored information and traditional discovery rules once again demonstrates that, as certain as death and taxes, are unintended consequences. This article proposes a discovery rule amendment as a different approach to e-discovery.

The lawsuit began as standard patent dispute litigation. Qualcomm sued Broadcom, a longtime rival in computer chip production, alleging infringement on various patents.<sup>3</sup> Broadcom’s defense alleged that Qualcomm participated in the Joint Video Team (JVT),<sup>4</sup> an organization whose members shared certain licensing rights<sup>5</sup> for patents. Broadcom contended that Qualcomm, as a member of the JVT, had agreed to share the right to use the patents in suit with other JVT members. Hence, Broadcom had the right to use the allegedly infringing patents due to waiver.<sup>6</sup> Consequently, Broadcom sought to discover evidence of Qualcomm’s participation in the JVT.

---

<sup>1</sup> Interview with Jennifer Isaac, Proj. Mgr., Qualcomm, Inc., in San Diego, Cal. (Jul. 7, 2011).

<sup>2</sup> Ms. Isaac was a witness for Qualcomm, Inc. in *Qualcomm, Inc. v. Broadcom Corp.* (*Qualcomm I*), No. 05cv1958-B (BLM), 2008 WL 66932 (S.D. Cal. Jan. 7, 2008).

<sup>3</sup> *Id.* at \*1.

<sup>4</sup> *Id.* (“JVT is the standards-setting body that created the H.264 standard, which was released in May 2003 and governs video coding.”).

<sup>5</sup> Licensing right is the right to use another’s unexpired patent. Usually, one has to pay the patent owner for this right. The JVT team members pooled their patents so all team members can use each other’s patents.

<sup>6</sup> Waiver is the sharing of licensing rights among a group of entities. Each participant contributes one or more patents, and every participant in the group will have the right to use the contributed patents, without having to pay licensing fee.

Ms. Isaac was an engineer working for Qualcomm at the time of the 2008 deposition.<sup>7</sup> In her deposition, Broadcom's counsel repeatedly asked her about her attendance at a JVT conference in 1993. At the time, Ms. Isaac did not remember or know which conference she attended fifteen years before. In fact, neither did her colleagues at Qualcomm, nor Qualcomm's counsel in the case. Unfortunately for Qualcomm, it did not find any email relating to its employees' participation in the JVT, because it did not search the right databases or use the right search terms. Later, when Adam Bier, Qualcomm's counsel, searched Viji Raveendran's<sup>8</sup> computer using the keyword "avc\_ce," he discovered an email chain sent to JVT participants.<sup>9</sup> This discovery undercut Qualcomm's denial of their participation in the JVT.

This simple but very late search began a sanction motion, in which Magistrate Judge Barbara Major characterized Qualcomm attorneys' failure to produce the emails as "actively organized" to conceal information from opposing counsel and the court.<sup>10</sup> In addition to a payment of \$9,259,985.09 for Broadcom's attorney's fees, Qualcomm and its counsel also faced the court's sanctions for discovery violation.<sup>11</sup> Qualcomm's counsel were the biggest losers in this suit, because they had unblemished records prior to the Broadcom suit,<sup>12</sup> but fell prey to the monsters of electronic discovery: search terms.<sup>13</sup>

Computers have become an essential part of life. Laptop computers, i-Pads and i-Phones all have the capacity to store vast amount of information. Most businesses now keep their records in electronic form. Just as computers are now common in everyday life, discovery involving computers and electronically stored information

---

<sup>7</sup> See *supra* note 2.

<sup>8</sup> Viji Raveendran was a witness and a Qualcomm employee.

<sup>9</sup> *Qualcomm I*, 2008 WL 66932, at \*3.

<sup>10</sup> *Id.* at \*5.

<sup>11</sup> *Id.*

<sup>12</sup> A search on Nov. 18, 2011 at <http://members.calbar.ca.gov/fal/MemberSearch/QuickSearch> showed all attorneys for Qualcomm referred to the State Bar of Cal. for sanction in the *Qualcomm I* case have clear records, due to the subsequent overturn of this case with regards to attorneys' sanctions. The attorneys were James R. Batchelder, Adam A. Bier, Kevin K. Leung, Christian E. Mammen, Lee Patch and Stanley Young. Lee Patch is currently an inactive member of the Cal. State Bar due to MCLE noncompliance. See also *Qualcomm, Inc. v. Broadcom Corp. (Qualcomm II)*, No. 05cv1958-RMB (BLM), 2008 WL 638108, at \*3 (S.D. Cal. Mar. 5, 2008).

<sup>13</sup> This case was partly overturned in *Qualcomm II*. The six attorneys listed above objected to their sanctions and consequently had their sanctions overturned. Qualcomm, however, did have to pay Broadcom's attorney fees.

(ESI) is now common in litigation. The Federal Rule of Civil Procedure (FRCP) addressed electronic discovery (e-discovery) for the first time in 2006.<sup>14</sup> The amended FRCP 26(f)(3)(C) states “A discovery plan must state that parties’ view and proposals on . . . any issues about disclosure or discovery of electronically stored information, including the form or forms in which it should be produced.”<sup>15</sup> The requirements for certification, including reasonable search, burden and expense, access and privilege, remain untouched in e-discovery. The parties’ responsibilities and rights remain the same in e-discovery, as in traditional discovery, despite the great difference between e-data and “hard copy” data.<sup>16</sup>

However, e-discovery bears a stark difference from discovery of other data forms, such as tangible form or “hard copy.” The volume of data available in electronic form is overwhelmingly large compared to anything an individual or an organization can store in paper. For example, a laptop computer has 120 gigabytes of memory.<sup>17</sup> One gigabyte is equivalent to 500,000 pages of documents.<sup>18</sup> One laptop alone can contain up to sixty million pages of documents. Manually examining this amount of information to find responsive documents is simply infeasible. Currently, to sift through this massive amount of data, attorneys have only one choice: utilize a search engine to search relevant databases to find responsive documents.

Consider a case with two parties involved in an employment dispute. The plaintiff, an aggrieved former employee, seeks discovery of five employees working for the defendant corporation. The discovery involves the five employees’ emails and other electronic records. The defendant’s attorneys must consolidate the relevant database, then use keyword search terms or other search methods to retrieve relevant emails and documents relating to the plaintiff’s discovery request. Another choice is to search each database individually. Choosing the right search method thus becomes the key to finding relevant documents and can sink the attorneys’ boat, if it somehow does not produce enough responsive documents. Finally,

---

<sup>14</sup> Bennett B. Borden et al., *Four Years Later: How the 2006 Amendments to the Federal Rules Have Reshaped the e-Discovery Landscape and Are Revitalizing the Civil Justice System*, 17 RICH. J.L. & TECH. 2 (2011).

<sup>15</sup> FED. R. CIV. P. 26 (f)(3)(C).

<sup>16</sup> Hard copy, in contemporary language, refers to information on paper.

<sup>17</sup> This capacity, like other features in computers, is likely to increase in a short time. Computers with larger hard drives are now rather common.

<sup>18</sup> *In re Seroquel Prods. Liab. Litig.*, 244 F.R.D. 650, 654 (M.D. Fla. 2007) (citing the MANUAL FOR COMPLEX LITIGATION (4th ed. 2004)).

after many searches and another manual review, the documents turned over are certified as being produced after “reasonable effort.” A reasonable search of electronic data is evaluated under the same standard as a reasonable search of tangible evidence. Attorneys are under the same scrutiny in e-discovery, as compared to discovery of tangible evidence, despite the large scope and difficulties unique to e-discovery.

This note proposes amendments to FRCP 26 to prevent disputes related to searching ESI in litigation discovery and reduce the cost and risks in e-discovery. It also explores the scope, cost, and technical difficulties of searching ESI in litigation, as well as various problems that attorneys and courts face in search terms disputes. Finally, it discusses the call for cooperation in e-discovery and different approaches courts and attorneys have tried, to resolve search problems in e-discovery. A mandated agreement on search method prior to e-discovery, coupled with a different standard for reasonable search in e-discovery, can prevent unnecessary disputes and reduce the cost of litigation.

## II. HOW A RULE AMENDMENT CAN HELP

As in the example above, search terms are currently the bait used to fish in the ocean of ESI. Search terms determine what will be retrieved, and ultimately, what will be produced. Attorneys and courts have been struggling to resolve this important issue and create an effective approach to e-discovery. Given the importance of search terms and search methods in retrieving responsive documents, a rule amendment, with guidance on how to search for ESI, is imperative.<sup>19</sup>

Currently, FRCP 26(f)(3)(C) requires parties to state in the discovery plan the proposals on “any issues about disclosure or discovery of electronically stored information . . .” While this clause allows parties to propose specific issues relating to searching ESI, it does not mandate it.<sup>20</sup> In fact, parties do not normally have a specific agreement on how to search for responsive documents in ESI. The

---

<sup>19</sup> FRCP 26 (f)(3)(C) governs electronic discovery. An amendment to this subpart is recommended.

<sup>20</sup> FRCP 26 (f)(3)(C) states “Discovery Plan. A discovery plan must state the parties’ views and proposals on . . . any issues about disclosure or discovery of electronically stored information, including the form or forms in which it should be produced.” Only parties’ views and proposals are required by the Rule, but not an agreement. Parties can choose to propose a plan on searching ESI, but are not mandated to do so.

Sedona Conferences' call for cooperation and early discussion reflects parties' lack of attention to search methodology and the consequences of this inattention.<sup>21</sup>

To prevent this problem, FRCP 26(f)(3)(C) should include language mandating that parties propose and agree on the specific methods each party will use to search their ESI. Presently, most of the disagreement over search methodology focuses on search term disputes.<sup>22</sup> With the pace of technological advancement, search term disputes may not last for long. To effectively address this issue, the FRCP addressing search method should be technologically adaptable.<sup>23</sup>

FRCP 26(f)(3)(C) should be amended to include a requirement that parties reach an agreement on specific search methods, as warranted by the parties' technological capacity, before conducting a search of their respective ESI. This agreement shall include the search method to be used, and the databases, storage spaces,<sup>24</sup> and equipment<sup>25</sup> where the search will be conducted. FRCP 26(f)(3)(C) should read "A discovery plan must state that party's view and proposals on, *and conclude with an agreement on . . . any issues about searching electronically stored information, including search method, data location to be searched, disclosure, or discovery of the same, including the form or forms in which it should be produced.*"<sup>26</sup>

---

<sup>21</sup> See William Butterfield et al., *The Case for Cooperation*, 10 SEDONA CONF. J. 339 (2009); see also Jason R. Baron et al., *The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in e-Discovery*, 8 SEDONA CONF. J. 189, 198 (2007).

<sup>22</sup> See RONALD J. HEDGES, DISCOVERY OF ELECTRONICALLY STORED INFORMATION – SURVEYING THE LEGAL LANDSCAPE 16 (2007) (citing Ad Hoc Comm. for Elec. Discovery of the U.S. Dist. Ct. for the Dist. of Del., *Default Standard for Discovery of Electronic Documents*, (2011), <http://www.ded.uscourts.gov/sites/default/files/Chambers/SLR/Misc/EDiscov.pdf>).

<sup>23</sup> See *infra* Sections III and IV. Search methods evolved from Boolean search method to natural language and concept search. What will be the future search method is open to speculation. A technologically adaptable rule will be flexible enough to account for unknown changes.

<sup>24</sup> Storage space, apart from hardware, can be virtual servers. Cloud computing is becoming increasingly popular and information may not be stored in any specific hardware in the future.

<sup>25</sup> The "equipment" is to account for any hardware a person may have, where ESI is stored.

<sup>26</sup> The italicized part is the proposed amendment. The added language mandates parties to reach an agreement on search methodology and locations in searching ESI during discovery.

This amendment serves two purposes: First, it mandates parties to discuss and reach an agreement on how to search before setting out to conduct the search. Second, it prevents disagreements in the later stage of litigation. Before the FRCP 26(f) conference, parties must craft a search plan, which forces parties to consider their search methods and the associated costs. An agreement will also minimize the cost and time involved in searching for responsive documents in ESI. With a known search method, parties do not need to experiment on their own to find an effective search method. Moreover, a search method agreement comes with an inherent requirement for parties to consider their systems and technical capabilities before agreeing to a certain search method. This would limit the chance of parties pleading undue burden in refusing to apply a previously agreed-upon search method.

Mandating an agreement on search method will help parties in orienting their search effort but it will not be enough to address the problem of under-producing due to an inefficient search method. Even with an agreement, the chance that parties will miss certain documents still exists.<sup>27</sup> In order to address this problem, the standard for electronic discovery under FRCP 26(g)(1)(A), needs to be specific to e-discovery.

An agreement on search method, even if mandated, will be of little force if such an agreement is not enough to serve the purpose of a “complete and correct disclosure.”<sup>28</sup> Under the current FRCP 26(g)(1)(A), by signing, an attorney or party certifies that “. . . [w]ith respect to a disclosure, it is complete and correct as of the time it is made.”<sup>29</sup> This standard, with regards to documents retrieved using search method, is somewhat like shooting a moving target. The target, or the “complete and correct disclosure,” is now an elusive goal. Just as a fisherman can never tell how many fish are left in the ocean, an attorney can never tell how many responsive documents are left in a client’s database.

Parties can only produce documents responsive to the search method and search terms applied. A “complete and correct disclosure” in e-discovery, even with the best effort, can only be complete and correct with regards to the search terms and search methods used and

---

<sup>27</sup> An agreement does nothing to change the efficacy of the search method employed.

<sup>28</sup> FED. R. CIV. P. 26 (g)(1)(A).

<sup>29</sup> See FED. R. CIV. P. 26 (g)(1)(A). There is no separate requirement for e-discovery certification. Attorneys are subject to the same standard under this Rule for all certifications in discovery production.

the databases searched. For production of ESI, the standard should be changed to “with respect to a disclosure, it is complete and correct, using the agreed upon search method, in relevant databases.” FRCP 26(g)(1)(A) should read “with respect to a disclosure, it is complete and correct as of the time it is made; *with respect to electronically stored information disclosure, it is complete and correct using the search method agreed upon at the 26(f) conference, in agreed upon storage spaces.*”

With these two amendments to FRCP 26(f)(3)(C) and FRCP 26(g)(1)(A), parties must consider and evaluate search methods before discovery.<sup>30</sup> Any disagreement can be resolved at the beginning and not at the end of the case. Furthermore, search method agreement provides a solid framework for searching and retrieving responsive documents. It is the equivalent of having bait before setting out to fish. The new standard of certification, suggested by an amendment to FRCP 26(g)(1)(A), is the equivalent of assuring the would-be fishermen that as long as the right bait (search methods) is used in the correct water (ESI databases), the “fish” (responsive documents) will be accepted as complete.

These amendments are necessary to limit risks in conducting e-discovery, while reducing the chance of discovery disputes. With known search methods and agreed upon search locations, attorneys will find e-discovery more manageable. This is more so, given the overwhelming scope of searching electronically stored information.

### III. THE SCOPE AND COST OF SEARCHING ELECTRONICALLY STORED INFORMATION

As above, a laptop computer can contain data equivalent to sixty million pages of documents. A corporation’s database is typically measured in terabytes, with one terabyte equivalent to one thousand gigabytes. A data processing center for a major corporation can contain ten thousand tapes or more.<sup>31</sup> With one tape equivalent to one

---

<sup>30</sup> The amendment to FRCP 26(f)(3)(C) mandates an agreement on search method, and the FRCP 26(g)(1)(A) allows certification according to the agreement under FRCP 26(f)(3)(C). To reach an agreement, a party must consider its capacity, research the possible locations of its data and acquaint itself to possible search methods.

<sup>31</sup> JAMES N. DERTOUZOS ET AL., RAND INST. FOR CIVIL JUSTICE, THE LEGAL AND ECONOMIC IMPLICATIONS OF ELECTRONIC DISCOVERY – OPTIONS FOR FUTURE RESEARCH 1 (2008), *available at* [http://www.rand.org/pubs/occasional\\_papers/2008/RAND\\_OP183.sum.pdf](http://www.rand.org/pubs/occasional_papers/2008/RAND_OP183.sum.pdf).

terabyte, converting ten thousand tapes into paper will result in a two-hundred-mile-high stack of paper.<sup>32</sup> Manually searching such a database is impractical. The scope of electronic discovery is therefore of unprecedented magnitude.

To search through this amount of data, attorneys currently utilize keyword searches. This is very similar to a Google search. A person keys in the “keyword(s),” sometimes with Boolean logic.<sup>33</sup> The search engine searches through the database and finds documents with keywords matching the search criteria (responsive documents). Responsive documents are displayed and the person retrieves those documents for further review. This process is deceptively simple, but it is not simple in practice. Google, one of the strongest search engines on Earth, provides the internet at one’s fingertips.<sup>34</sup> An individual’s or an organization’s database must be searched without Google’s search engine or, sometimes, any kind of search engine at all.<sup>35</sup>

Not all data stored in a computer is the same. Metadata<sup>36</sup> is probably the most difficult data to search. Metadata contains “data about data,” including the time of creation, author, comments, deleted keystrokes, volume, and data location on the hard drive.<sup>37</sup> The wealth of information obtained from metadata can be the key to a party’s case.<sup>38</sup> While simple keyword searches usually produce responsive documents, metadata is not part of the documents’ content and does not respond to keyword searches.<sup>39</sup>

---

<sup>32</sup> *Id.*

<sup>33</sup> David H. Tennant et al., *Best Practices in E-discovery in New York State and Federal Courts*, 2011 N.Y. BAR ASSOC. SEC. COMM’L & FED. LITIG. REP. 30, (July 2011), available at <http://www.nysba.org/AM/Template.cfm?Section=Home&ContentID=56437&Template=/CM/ContentDisplay.cfm> (Boolean search logic provides search algorithms such as “and,” “or,” “in the same sentence,” “in the same paragraph,” “within *x* words of,” “not.”).

<sup>34</sup> Paul Gil, *The 10 Best Search Engines of 2011*, ABOUT.COM (Apr. 2012), [http://netforbeginners.about.com/od/navigatingthenet/tp/top\\_10\\_search\\_engines\\_for\\_beginners.htm](http://netforbeginners.about.com/od/navigatingthenet/tp/top_10_search_engines_for_beginners.htm) (last visited Oct. 31, 2011).

<sup>35</sup> WINDOWS 7 has the capacity to search documents in each drive using Boolean search terms. However, the process is slow, with unknown efficacy.

<sup>36</sup> SHIRA A. SCHEIDLIN, MOORE’S FED. PRAC., E-DISCOVERY: THE NEWLY AMENDED FEDERAL RULES OF CIVIL PROCEDURE 43 (2006 Special Pamphlet), available at LexisNexis (Metadata is information describing the history, tracking, or management of an electronic file); BRENT D. ROPER, USING COMPUTERS IN THE LAW OFFICE 559 (5th ed. 2008); Tennant, *supra* note 33, at 36.

<sup>37</sup> ROPER, *supra* note 36.

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

<sup>39</sup> *Id.*

Metadata, despite its retrieval difficulties, can be the linchpin in a case's e-discovery and search of ESI. In one case, a woman sued her former employer, alleging age discrimination in her layoff.<sup>40</sup> She requested to discover the spreadsheet's native format<sup>41</sup> that her former employer used to analyze layoff options.<sup>42</sup> The employer provided the file without metadata, by locking<sup>43</sup> certain cells and formulas.<sup>44</sup> Metadata was the only way to discover the former employee's requested information and the court ruled that the employer must produce the file in the native format.<sup>45</sup> This case illustrates the importance of metadata in e-discovery.

Once a file is "deleted" from a computer, it goes into the "Recycle Bin." This only means the file is marked as "deleted," freeing the space for future use, but real data still exists, which is why a file can still be "recovered" from the Recycle Bin.<sup>46</sup> Even when a file is deleted from the Recycle Bin, data from that file still exists, until the space is overwritten by new data.<sup>47</sup> This process is called "soft deletion" and soft deletion can be recovered with complete integrity.<sup>48</sup> It is similar to a fingerprint on a table. Even when one removes her finger from the table, a fingerprint stays behind and can be retrieved using forensic methods. Retrieval of metadata is difficult, as a special program to access such data is needed. The form of production may

---

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

<sup>41</sup> *Id.* at 557 (Native format is the file structure as defined when it was first created. For example, a Word document created by Word 2010 has the native format as .docx).

<sup>42</sup> SCHEIDLIN, *supra* note 36, at 43; ROPER, *supra* note 36, at 43; Tennant, *supra* note 33, at 36.

<sup>43</sup> "Locking" a cell in Excel is performing a lock function to a cell, such that only the final value of computation is shown, but not the formulae underlying the computation. If cell C1's formulae is  $C1 = A1+B1$ , and A1 has a value of 3, B1 has a value of 4, then C1 will have a final value of  $3+4=7$ . When cell C1 is locked, one can only see the number 7 when clicking on cell C1, and not the formulae " $C1 = A1+B1$ ."

<sup>44</sup> ROPER, *supra* note 36.

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

<sup>46</sup> SHIRA A. SCHEINDLIN ET AL., THE SEDONA CONFERENCE, ELECTRONIC DISCOVERY AND DIGITAL EVIDENCE: CASES AND MATERIALS 9 (2009).

<sup>47</sup> Another popular action conducted by IT personnel before deploying an old hard drive a new user is "formatting," which eliminates all electronic content. In fact, data still exists and is recoverable with the help of a computer forensic specialist. See MICHELE C.S. LANGE & KRISTIN NIMSGER, ELECTRONIC EVIDENCE AND DISCOVERY: WHAT EVERY LAWYER SHOULD KNOW NOW 221 (2d ed. 2009).

<sup>48</sup> ROPER, *supra* note 36, at 563 (stating that complete integrity is the original format of data before deletion).

also be an obstacle, because a party may not have the software or expertise necessary to access metadata in its native form.

With recent advances in technology, computers' capacity has increased rapidly. Together with increased capacity, the cost to store information electronically plummets. In 1990, the cost to store one gigabyte of data was \$20,000, as compared to less than \$1 in 2007.<sup>49</sup> In contrast, the cost to review documents remains approximately the same, as document review still relies on human effort. If one assumes a billing rate of \$200 per hour for a junior associate, the cost of reviewing one gigabyte is over \$30,000.<sup>50</sup> The difference in the cost of storing as compared to the cost of reviewing, exemplifies the burden and cost in e-discovery.<sup>51</sup>

The cost of retrieving computer data can be staggering, as well. Recovering metadata from one computer hard drive can cost \$51,000,<sup>52</sup> since a computer forensic specialist must be involved in the retrieval.<sup>53</sup> The cost of searching, storing, and reviewing responsive documents is not included. In a recent case of a simple employment dispute, the plaintiff requested metadata production from the defendant. The court found the cost of searching metadata outweighed any potential benefit of finding responsive documents and denied the plaintiff's request for such a search.<sup>54</sup>

Even a search through an employee's email over a period of two years, using keywords related to a specific patent can cost "tens of thousands of dollars for initial search," not including attorney time.<sup>55</sup> In *Eurand v. Mylan*, both parties disputed over search terms to be used.<sup>56</sup> The case was a patent dispute, involving patent "793"<sup>57</sup> and

---

<sup>49</sup> Michelle Kessler, *Day of Officially Drowning in Data Almost Upon Us*, USA TODAY (Mar. 7, 2007), [http://www.usatoday.com/tech/news/2007-03-05-data\\_N.htm](http://www.usatoday.com/tech/news/2007-03-05-data_N.htm).

<sup>50</sup> Jason R. Baron et al., *The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in e-Discovery*, 8 SEDONA CONF. J. 189, 198 (2007).

<sup>51</sup> In recent years, electronic document reviewing has become more popular. However, having a pair of eyes to look over the documents before production is still standard operation in most cases.

<sup>52</sup> See *Jacobeit v. Rich Township High Sch. Dist. 227*, No. 1:09-cv-01924, 2011 WL 2039588, at \*8 (Ind. Ill. May 25, 2011).

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

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

<sup>55</sup> *Eurand, Inc. v. Mylan Pharm.*, 266 F.R.D. 79, 84 (D. Del. 2010).

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

<sup>57</sup> Patents are referred to by the last three digits in their patent numbers. A patent numbered 125,678,725 is thus referred to as patent '725.

alleged inequitable conduct.<sup>58</sup> Mylan Pharmaceuticals, Inc. sought to discover what the inventors knew and were saying about prior arts, during the prosecution of the “793” patent.<sup>59</sup> Mylan suggested limited search terms, including terms such as “the 215 patent,” “the methylphenidate patent,” “Razaghi,” and “amitriptyline with extended release.”<sup>60</sup> Eurand claimed that, even with these limited search terms, the search still cost tens of thousands of dollars.<sup>61</sup> Eurand also contended that the proposed search terms had nothing to do with Mylan’s inequitable conduct claim, the claim for which Mylan sought discovery.<sup>62</sup> The cost of searching was used as a defense against the discovery request.<sup>63</sup>

The cost of searching is a burden, not only for litigating parties, but also for third parties, whose are compelled by *subpoena duces tecum* to produce documents.<sup>64</sup> Fannie Mae<sup>65</sup> was the defendant in a securities lawsuit, concerning its alleged deceptive accounting practices. As the Office of Federal Housing Enterprise Oversight (OFHEO) oversaw Fannie Mae, it received a court order to produce documents related to its investigation into Fannie Mae’s alleged deceptive practice.<sup>66</sup> The defendants in the suit, former executives of Fannie Mae, sought to discover the relevant investigative documents to prove their transparency with OFHEO.<sup>67</sup> The subpoena order included a list of search terms totaling four hundred keywords, aimed at producing six hundred thousand documents.<sup>68</sup> OFHEO initially moved to quash the subpoena, but the court ruled against OFHEO. OFHEO was forced to hire fifty contract attorneys to review the documents in an effort to comply with the court order.<sup>69</sup> The cost of the review eventually reached over six million dollars, which was more than nine

---

<sup>58</sup> *Eurand, Inc.*, 266 F.R.D. at 84.

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

<sup>60</sup> *Id.* at 84 n.23.

<sup>61</sup> *Id.* at 84.

<sup>62</sup> *Id.*

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

<sup>64</sup> FED. R. CIV. P. 45.

<sup>65</sup> *Company Overview*, FANNIE MAE, <http://www.fanniemae.com/portal/about-us/company-overview/about-fm.html>? (last visited Apr. 13, 2012) (Fannie Mae is a government-sponsored enterprise, aiming at expanding the secondary mortgage market by securitizing mortgages in the form of mortgage-backed securities.).

<sup>66</sup> *In re Fannie Mae Sec. Litig.*, 552 F.3d 814, 817 (2009).

<sup>67</sup> *Id.*

<sup>68</sup> *Id.*

<sup>69</sup> *Id.*

percent of the agency's entire annual budget.<sup>70</sup> Perhaps due to the exorbitant cost, the agency could not meet the deadline and the defendants filed a motion to hold OFHEO in contempt of the court.<sup>71</sup> The motion was granted with additional sanctions against OFHEO.<sup>72</sup>

The above cases show the burden of e-discovery, particularly from the document searching perspective, in civil litigation. This burden inflicts difficulties to all types of parties, from private individuals to sophisticated corporations and even to governmental entities. Litigation does not need to be overly complicated and e-discovery should not be an added obstacle to justice.

#### **IV. TECHNICAL DIFFICULTIES AND OTHER PROBLEMS IN USING KEYWORD SEARCH TERMS TO FIND RESPONSIVE DOCUMENTS**

Searching ESI using keywords can be a simple task. Most lawyers perform this search using Westlaw or LexisNexis, popular legal databases available in electronic form.<sup>73</sup> In addition to typing in the keywords, one can use Boolean search terms to more accurately describe the terms needed. Simple Boolean search terms include "&" (and), "!" (anything), "-" (but not), or more complicated terms, such as "/p" (in the same paragraph) or "/s" (in the same sentence). However, the capability to interpret and search using Boolean search terms may not be present in every database. Some programs or databases may have different search capabilities. For example, searching the text in one document for a phrase may be available in Adobe Reader, but may not be available in a Notepad text file.<sup>74</sup>

Data organization in a party's possession may also pose a problem. An organization has multiple computers, and its data is stored on different servers, different hard drives or different databases.<sup>75</sup> Even searching with one search term, one must search each computer and each database separately. For instance, the email server of the Rich

---

<sup>70</sup> *Id.*

<sup>71</sup> *Id.* at 818.

<sup>72</sup> *In re Fannie Mae*, 552 F.3d at 818.

<sup>73</sup> SCHEINDLIN et al., *supra* note 36.

<sup>74</sup> *How to Use Adobe Search Function*, VILLAGEOFLENCOE.ORG, [http://www.villageofglencoe.org/assets/1/finance/HOW\\_TO\\_USE\\_THE\\_ADOBE\\_SEARCH\\_FUNCTION.pdf](http://www.villageofglencoe.org/assets/1/finance/HOW_TO_USE_THE_ADOBE_SEARCH_FUNCTION.pdf) (last visited Oct 20, 2011).

<sup>75</sup> Searching on a company's server is a must-do, but local hard drives, CD-ROM, and thumb drives should also be searched. See SHARON D. NELSON ET AL., *THE ELECTRONIC EVIDENCE AND DISCOVERY HANDBOOK FORMS, CHECKLISTS, AND GUIDELINES 2* (2006).

Township School District could not be searched as a whole, but had to be searched separately for each employee's email inbox.<sup>76</sup> For a large organization, searching each computer separately poses a huge challenge, especially in identifying the custodians whose computers must be searched. In the case discussed in the introduction, Qualcomm's failure to search one computer with one particular keyword resulted in massive discovery sanctions.<sup>77</sup> Data organization may undoubtedly increase the burden on the producing party.

Attorneys' familiarity with information technology poses another problem for e-discovery in general and for determining search terms and search method in particular. Most attorneys do not receive information technology training as part of their curriculum in law school or their previous education.<sup>78</sup> Meanwhile, in e-discovery, attorneys have to deal with nuance details about how electronic information is stored, both permanently and temporarily.<sup>79</sup> With the FRCP 2006 Amendment and the prevalence of ESI in litigation, courts no longer tolerate attorneys who do not understand e-discovery.<sup>80</sup> In fact, litigation attorneys are forced to understand basic information technology or seek expert help.<sup>81</sup>

## V. KEYWORD SEARCHING IS REALLY A PROBLEM AND NOT A SOLUTION

Efficiency is a major problem in using keyword search to retrieve documents. The reality is, even with the best search engine, the only

---

<sup>76</sup> *Jacobbeit*, 2011 WL 2039588, at \*7.

<sup>77</sup> See *Qualcomm I*, No. 05cv1958-B (BLM), 2008 WL 66932 (S.D. Cal. Jan. 7, 2008).

<sup>78</sup> Law schools now have Westlaw and LexisNexis training. However, law schools do not provide a course specializing in information technology.

<sup>79</sup> See JOHN M. BARKETT, *RAM and Cache? Things You Did Not Learn in Law School*, in *E-DISCOVERY TWENTY QUESTIONS AND ANSWERS* 7 (2008).

<sup>80</sup> See *U.S. ex rel. Tyson v. Amerigroup Ill., Inc.*, No. 02-C-6074, 2005WL 3111972, at \*3 (N.D. Ill. Oct. 21, 2005) ("Significantly, the defendants do not challenge any of Mr. Perry's assertions. This omission, they claimed at oral argument, was the inevitable result of having no familiarity with the internal systems used at HFS. The argument is unpersuasive. The defendants could have sought leave to depose Mr. Perry, and, of course, they could have retained an expert of their own to opine on the validity of Mr. Perry's statements—at least in a general sense.") The court in this case rejected an argument that the defendant's lack of understand of the plaintiff's information technology system is an excuse.

<sup>81</sup> Geoffrey A. Vance & Courtney I. Barton, *Drowning in Zubulake: The Rules, Pitfalls, and Benefits of Electronic Discovery*, 10 NAT'L LEGAL CTR. FOR THE PUB. INTEREST 1, 30 (2006), <http://www.mwe.com/info/pubs/gvance0406.pdf>.

documents retrieved are those that satisfy the keywords criteria. A search using the keyword “Bronx” results in all documents containing this word, regardless of context.<sup>82</sup> A construction company working on two projects, Bronx Zoo and Bronx County Hall of Justice will have emails relating to both projects. If litigation only concerns the Hall of Justice project, applying the keyword “Bronx” results in production of documents relating to the Zoo project as well. Whether the keywords criteria are effective in retrieving documents is a guess at best.

Research has shown that even a team of experienced lawyers and paralegals could only find twenty five percent of relevant documents using keyword search.<sup>83</sup> As compared to the team’s expectation of seventy five percent, this is a dismal result.<sup>84</sup> In litigation, missing out on seventy five percent of relevant data dramatically changes the parties’ knowledge and position in the case. This research was conducted in 1985 and search techniques certainly were not as efficient back then.<sup>85</sup> One would guess that search efficacy has improved in recent years. However, the Sedona Conference once again affirmed that efficacy in using keyword search method still had not been measured.<sup>86</sup> Uncertain efficacy is thus a major problem in keyword search method.<sup>87</sup>

Determining and designing keywords are recognized problems since keywords determine which documents are retrieved. Designing keywords requires input from data custodians, parties, computer experts, and attorneys. While keywords are crucial to searching ESI, neither the FRCP nor the courts have given clear guidelines. Testing the reliability of the search is another important aspect of e-discovery. However, the law in this area is limited and depends entirely on the jurisdiction.

---

<sup>82</sup> See *William A. Gross Const. Assoc., Inc. v. Am. Mfrs. Mut. Ins. Co. (Gross)*, 256 F.R.D. 134 (S.D.N.Y. 2009) (One of the proposed search terms was “Bronx! but not Zoo” because Hill, the producing party, was working on another project called Bronx Zoo. Searching for “Bronx” necessarily will result in numerous hits about the Bronx Zoo project, which is unrelated to the case.).

<sup>83</sup> David Blair & M.E. Maron, *An Evaluation of Retrieval Effectiveness For A Full-Text Document Retrieval System 3* (Graduate Sch. of Bus. Admin. Univ. of Mich. Res., Working Paper No. 364, 1984).

<sup>84</sup> *Id.* at 12.

<sup>85</sup> Blair, *supra* note 83.

<sup>86</sup> Baron, *supra* note 50, at 197.

<sup>87</sup> See *Qualcomm I*, No. 05cv1958-B (BLM), 2008 WL 66932 (S.D. Cal. Jan. 7, 2008) (the efficacy or the lack thereof in using keyword search is well demonstrated in the case illustrated at the beginning of this article).

In *Stanley v. Creative Pipe*,<sup>88</sup> the plaintiff sought a ruling that the documents the defendant produced were not within the attorney-client privilege and work product doctrine. Creative Pipe contended that the documents produced were among the nontext-searchable ESI.<sup>89</sup> Due to time constraint, Creative Pipe's attorneys only reviewed by glancing through the title pages of nontext-searchable ESI documents.<sup>90</sup> Consequently, Creative Pipe asserted that these documents were inadvertently produced.<sup>91</sup> The court, however, found that Creative Pipe did not show that the documents were among the nontext-searchable ESI.<sup>92</sup>

Specifically, Magistrate Judge Paul Grimm questioned how the keywords were developed, how the search was conducted and what quality controls were employed to assess their reliability,<sup>93</sup> and even the qualification of the attorneys and parties, who designed the keywords.<sup>94</sup> The court also emphasized that parties must test the efficacy of the search method after conducting the search.<sup>95</sup> This case is an example of how keyword search can result in inadvertent production of privileged documents<sup>96</sup> and the difficulties in retracting them once they are produced.

Furthermore, keyword searches work best with documents containing predictable language, where the keywords have the same meaning, regardless of context.<sup>97</sup> Yet often, this is not the case. For example, in an employment law case, where a former employee alleges age discrimination in a layoff, the keywords used may include "layoff," "fired," "let go," etc., but if the company's personnel use their own code words in emails, it would be almost impossible to guess what those words should be. False positives or documents containing the keyword(s), but have no relation to the case are huge

---

<sup>88</sup> *Id.*

<sup>89</sup> *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 250 F.R.D. 251, 256 (D. Md. 2008).

<sup>90</sup> *Id.*

<sup>91</sup> *Id.*

<sup>92</sup> *Id.*

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

<sup>94</sup> *Id.*

<sup>95</sup> *Victor Stanley, Inc.*, 250 F.R.D. at 257 (The only prudent way to test the reliability of the keyword search is to perform some appropriate sampling of the documents determined to be privileged.).

<sup>96</sup> *Id.* at 256–57 (“[a]ll keywords searches are not created equal; and there is a growing body of literature that highlights the risks associated with conducting an unreliable or inadequate keyword search or relying exclusively on such searches for privilege review.”).

<sup>97</sup> Baron, *supra* note 50, at 201.

distractions and waste attorneys' time.<sup>98</sup> In the example above, a false positive would be an email from one employee to another talking about how she "let go" of her dreams. This email has no relation to the case at hand.

Another issue in documents created on computers is misspelling. While some software includes a spell check function, not all software does.<sup>99</sup> Consequently, misspelled words will not be "responsive" to the keywords and those documents will not be retrieved.<sup>100</sup> A search using the keyword "striking" will not identify a document with "striking," despite the document's relevance.

One word can also have many variations from one "stem."<sup>101</sup> In searching for emails relating to a layoff, apart from searching for "fire," it is also necessary to search for "firing" and "fired."<sup>102</sup> Other words may have more variations, as the computer only retrieves documents that match exactly the search criteria, attorneys must think of all possible "stems" and all possible misspellings to increase the search's efficiency. If the search engine has Boolean search capacity, then one can search for "fire" and its variations using the search term "fir!," with the exclamation mark indicating "anything." For other words such as "drink," one must search for "drank," "drinking" and "drunk". There is not a single way to put these words in Boolean logic.

Normally, litigating parties must hire a vendor to extract data from their hardware and supply the vendor with search terms to conduct the search,<sup>103</sup> because vendors are involved in the harvesting, filtering, and production of data. The line between the attorney and the vendor must be clearly drawn. While the attorney chooses what to produce, the vendor, through his search, determines what is retrieved. The vendor's work, if not checked and monitored closely, can expose both client and attorney to sanctions.<sup>104</sup> As a result, attorneys must manage the vendors and their work closely.

---

<sup>98</sup> *Id.*

<sup>99</sup> MICROSOFT OFFICE has a spell check function. However, this function can be turned off. Some email programs have spell check functions, but not all of them do.

<sup>100</sup> Baron, *supra* note 50, at 202.

<sup>101</sup> *Id.*

<sup>102</sup> *Id.*

<sup>103</sup> Storing and searching e-data has become a booming business. The author visited one such vendor, where most of the business space was dedicated to hardware. The electricity bill to maintain this hardware collection and provide cooling for the storage space amounted to approximately \$20,000 a month.

<sup>104</sup> ROPER, *supra* note 36, at 565.

## VI. THE COURT'S AWKWARD POSITION IN THE SEARCH TERMS WAR

Like other discovery disputes, parties bring their search term problems to court when they cannot reach an agreement. The court must either mandate search terms or mandate cooperation. In the former, the court must act as a fact finder, a computer expert, and both sides' counsel, all at the same time.<sup>105</sup> Understandably, the court, with its already heavy burden and docket, laments this new responsibility.<sup>106</sup> In the latter, an order to cooperate does little to guide the parties on what exactly they should do. Parties need to know how to craft search terms. They also need to know how many search terms are enough. A third option, appointing a technical master to determine search terms and search method, has been used in certain situations.<sup>107</sup> This option has limitations, which will be discussed later in section VIII.

In *Gross*, after lamenting the court's awkward position and the lawyers' ignorance in designing search terms, the court provided search term guidelines and ordered cooperation between the parties.<sup>108</sup> In that case, the dispute involved the Bronx Criminal Court Complex construction.<sup>109</sup> The Dormitory Authority of the State of New York (DASNY) was the owner of the project<sup>110</sup> and Hill International was DASNY's construction manager.<sup>111</sup> The parties and Hill agreed to search Hill's email database.<sup>112</sup> DASNY proposed only a few search terms: "DASNY," "Dormitory Authority," "Authority," "Court! in connection with Bronx," "Hall of Justice" and "Bronx but not Zoo."<sup>113</sup> The other parties requested thousands of additional search terms,

---

<sup>105</sup> See, e.g., *United States v. O'Keefe*, 537 F. Supp. 2d 14, 24 (D.D.C. 2008).

<sup>106</sup> See *id.* ("Given this complexity, for lawyers and judges to dare opine that a certain search term or terms would be more likely to produce information than the terms that were used is truly to go where angels fear to tread.").

<sup>107</sup> See, e.g., *Medtronic Sofamor Danek, Inc. v. Michelson*, 229 F.R.D. 550 (W.D. Tenn. 2003); *In re Seroquel Prods. Liab. Litig.*, No. 6:06-md-1769-Orl-22DAB, 2008 WL 1995058 (M.D. Fla. 2008); see also *United States v. Philip Morris*, 449 F.Supp.2d 1 (D.D.C. 2006).

<sup>108</sup> *Gross*, 256 F.R.D. at 136.

<sup>109</sup> *Id.* at 134.

<sup>110</sup> *Id.*

<sup>111</sup> *Id.*

<sup>112</sup> *Id.*

<sup>113</sup> *Id.* (Hill worked on another project, Bronx Zoo. The search term "Bronx but not Zoo" will eliminate any emails relating to the Bronx Zoo project, which is not involved in the case.).

which essentially resulted in Hill's entire email database being produced.<sup>114</sup> Hill, the custodian of the emails, did not contribute any idea to the search terms list.<sup>115</sup> The court concluded "that left the court in the uncomfortable position of having to craft a keyword search methodology for the parties, without adequate information from the parties (and Hill)."<sup>116</sup>

In another search term dispute, *Eurand v. Mylan*, the court characterized keyword search techniques, as "mysteries."<sup>117</sup> Mylan Pharmaceuticals claimed the plaintiff had failed to apply certain search terms designed to reveal the inventors' thought and knowledge.<sup>118</sup> The plaintiff responded with a claim of undue burden of cost ("tens of thousands of dollars").<sup>119</sup> The court saw nothing but "mysteries . . . involved the interplay of computer technology, statistics and linguistics-complex."<sup>120</sup> Eventually, the court resorted to a reasonableness standard to determine the adequacy of the proposed search terms.<sup>121</sup> Such difficulties can be avoided, if the parties are mandated to reach an agreement on search method and the "reasonable effort" standard in e-discovery can be satisfied by compliance to the agreed upon search method.

## VII. THE CALL FOR COOPERATION

The call for cooperation in e-discovery has permeated through the legal profession.<sup>122</sup> While the FRCP does not address search methodology or mandate search method agreement, attorneys have no recourse other than cooperation in designing search terms. Once a party signs his name to a response to a discovery request, certifying that the production is, to the best of his knowledge after a thorough

---

<sup>114</sup> *Gross*, 256 F.R.D. at 134.

<sup>115</sup> *Id.*

<sup>116</sup> *See id.* at 135 (The Court went on in the discussion "This case is just the latest example of lawyers designing keyword searches in the dark, by the seat of the pants, without adequate (indeed, here, apparently without any) discussion with those who wrote the emails." The Court concluded with "Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords. . . It is time that the Bar – even those lawyers who did not come of age in the computer era – understand this.").

<sup>117</sup> *Eurand, Inc. v. Mylan Pharm.*, 266 F.R.D. 79, 84 (D. Del. 2010).

<sup>118</sup> *Id.* at 84.

<sup>119</sup> *Id.*

<sup>120</sup> *Id.*

<sup>121</sup> *Id.* at 85.

<sup>122</sup> *See supra* note 82.

search, he is at the mercy of his search method. Whether the opposing party can find other information using different search methods is left to fate. Cooperation minimizes the risk of under producing, once a party has agreed to a certain search method proposed by the opposing counsel, that party will be less likely to come back and claim inequitable conduct.

The Sedona Conference leads the call for cooperation in e-discovery, issuing its Cooperation Proclamation in 2008.<sup>123</sup> The following year, the Sedona Conference followed up with a related issue, *The Case for Cooperation*.<sup>124</sup> Cooperation starts with evaluating the discovery request and thinking twice before objection.<sup>125</sup> The scope of e-discovery, as outlined above, is immense and careless or strategically burdensome requests can overwhelm all parties.<sup>126</sup> An overwhelmed party ultimately will resort to the court, wasting judicial, as well as, parties' resources on meaningless discovery disputes. On the other hand, boiler plate objections to discovery requests only work to frustrate an opposing party and the court.<sup>127</sup> Parties should at least spend the effort to study the discovery request carefully before sending it out or determine the most effective objections.<sup>128</sup> Otherwise, disputes may result in sanctions and simultaneously reduce the chance of adjudicating the case on the merits.<sup>129</sup>

The simplest reason for cooperation is economic, since strategic cooperation reduces the chance of disputes.<sup>130</sup> Each dispute results in costly discourses between the parties before the filing and arguing of motions. Courts then must spend time and resources to resolve each dispute. From the attorneys' standpoint, disputes create more work. From the parties' standpoint, disputes waste money and time.<sup>131</sup> From

---

<sup>123</sup> The Sedona Conference, *The Sedona Conference Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009).

<sup>124</sup> See William Butterfield et al., *The Case for Cooperation*, 10 SEDONA CONF. J. 339 (2009); see also Baron, *supra* note 50, at 198.

<sup>125</sup> The Sedona Conference, *supra* note 123, at 333.

<sup>126</sup> Butterfield, *supra* note 124.

<sup>127</sup> Butterfield, *supra* note 124.

<sup>128</sup> Butterfield, *supra* note 124.

<sup>129</sup> Butterfield, *supra* note 124.

<sup>130</sup> Butterfield, *supra* note 124 at 356 (“Strategic cooperation, rather than “dueling,” can reduce the time and cost in e-discovery. The most straightforward reason for parties to cooperate is economic: unnecessary, combative discovery wastes time and money.”).

<sup>131</sup> With billing rate for a junior associate at approximately \$200 an hour, a discovery motion, which will take tens of hours of work, can be very costly for both parties.

the court's standpoint, disputes waste time that could have been used to adjudicate a matter on the merits.<sup>132</sup> Given the duty of diligence to clients, attorneys ought to cooperate to reduce cost and time. Furthermore, as officers of the court and counselors at law, attorneys should be mindful of the time and economic effect of adversarial lawyering on both clients and the court. Effectively serving the client by cooperation will increase the public's trust in the law, the litigation process, and the legal profession, as a whole.

Parties are required to meet, by FRCP 26(f), before the start of discovery.<sup>133</sup> In this meeting, agreement on what search terms or search method will be used and how to craft search terms should be reached. Search terms are currently the keys to uncovering responsive data and documents in e-discovery. Without an agreement on search terms, parties are likely to craft search terms "in the dark, by the seat of the pants."<sup>134</sup> More importantly, should the search turn out to be inadequate the parties will have nowhere to turn and may face sanctions.<sup>135</sup> The Sedona Conference recommends talking early and cooperatively.<sup>136</sup> Discovery disputes should be addressed early in the case, instead of at the end, as the end of the case should be resolution on the merits.<sup>137</sup>

## VIII. OTHER SOLUTIONS ATTEMPTED BY THE COURT AND ATTORNEYS

### A. Technical Master

A technical master is a person appointed by the court, sometimes with the parties' agreement, to manage e-discovery and arbitrate technical disputes.<sup>138</sup> The exact duty of a technical master depends

---

<sup>132</sup> Ralph C. Losey, *Mancia v. Mayflower Begins a Pilgrimage to the New World of Cooperation*, 10 SEDONA CONF. J. 377, 389 (2009).

<sup>133</sup> FED. R. CIV. P. 26(f) ("(1) Conference timing. Except in a proceeding exempted from initial disclosure under Rule 26 (a)(1)(B) or when the court orders otherwise, the parties must confer as soon as practicable – and in any event at least 21 days before a scheduling conference is to be held or a scheduling order is due under Rule 16(b).").

<sup>134</sup> *Gross*, 256 F.R.D. at 135.

<sup>135</sup> *Qualcomm I*, No. 05cv1958-B (BLM), 2008 WL 66932 (S.D. Cal. Jan. 7, 2008).

<sup>136</sup> The Sedona Conference, *supra* note 123 at 332.

<sup>137</sup> Butterfield, *supra* note 124, at 343–44.

<sup>138</sup> Shira A. Scheindlin & Jonathan M. Redgrave, *Special Masters and E-discovery: The Intersection of Two Recent Revisions to the Federal Rules of Civil Procedure*, 30 CARDOZO L. REV. 347, 374 (2008).

largely on the case at hand. Where there is a dispute on search terms and search method, the technical master typically will make decisions regarding search terms, oversee the design of searches, and advise both parties on technical issues and other matters.<sup>139</sup> In a few cases in which complex e-discovery overwhelms the parties (and probably the court), a technical master can help resolve certain issues as outlined above. However, appointing a technical master can create other problems.

First, courts prefer to appoint a technical master who has received approval from both parties.<sup>140</sup> In the most likely scenario, the court must decide on the technical master, as parties are already arguing over the e-discovery dispute. For example, the court in *In re Seroquel Products Liability Litigation*,<sup>141</sup> appointed a technical master who was not recommended by either party, because both parties moved for their own technical master. Essentially, the court had to resolve an additional dispute: the appointment of the technical master. The added time and cost to appoint a technical master makes this solution viable only in complicated and high stake cases.

Second, a technical master has a strong financial incentive to “drag out” the dispute, instead of resolve it. A technical master will be paid for his service, most likely by the hour.<sup>142</sup> When the dispute resolves, the technical master’s work and payment end. While attorneys must abide ethics rules and put the client’s interest first; technical masters are not attorneys and are not bound by any rule, except for the court’s authority. In short, the technical master can actually be a problem, if acting in bad faith, as opposed to a solution in search term and e-discovery disputes.

Finally, and most importantly, the cost of a technical master is high. Parties must pay for a technical master from their own litigation fund. Unsurprisingly, courts appoint technical masters in limited cases, where e-discovery is complex or only after an apparent failure in e-discovery.<sup>143</sup> While the appointment of a technical master may limit

---

<sup>139</sup> *Medtronic*, 229 F.R.D. at 559.

<sup>140</sup> See *In re Seroquel Prods. Liab. Litig.*, No. 6:06-md-1769-Orl-22DAB, 2008 WL 1995058, at \*1 (M.D. Fla. May 7, 2008).

<sup>141</sup> *Id.*

<sup>142</sup> See *Chaset v. The Upper Deck Co.*, No. 830257-9, 2002 WL 33966731, at \*1 (Cal. Super. Ct. Nov. 7, 2002) (The court ordered the appointment of a technical master and split the cost between the parties in one decision.)

<sup>143</sup> See *id.* at 2 (special master ordered due to the amount of ESI at issue); *In re Seroquel Prods. Liab. Litig.*, No. 244 F.R.D. 650, 661 (M.D. Fla. May 7, 2008) (special master ordered after the court sanctioned the defendant for “purposeful

disputes, the price tag makes it impractical to implement in many cases.<sup>144</sup>

The economic burden of a technical master could be another obstacle for unsophisticated plaintiffs. A former employee seeking to discover electronic records of her employment from a former employer will likely be unable to afford a technical master. Civil rights organizations suing large governmental agencies also face a long e-discovery process of vast amounts of ESI. Appointing a technical master in such cases imposes huge up front litigation cost. For unsophisticated and non-profit parties, the technical master is an expensive obstacle pushing justice even farther away.

### **B. Other Search Methods**

Apart from appointing a technical master, courts have considered different approaches to search electronic documents. Using Boolean search terms is not the only way to retrieve responsive documents. Other search methods have been used, both in common search engines and in legal applications. Westlaw Next, the new Westlaw application, now uses natural language search, which is a search based on fuzzy logic.<sup>145</sup> Fuzzy logic aims at producing documents based on a linguistic understanding and not mechanical search criteria like Boolean logic. Thus, in Westlaw Next, searching for a word with a particular meaning can return a document having a different word with the same meaning.<sup>146</sup>

---

sluggishness” in its production of ESI); *see also* United States v. Philip Morris, 449 F.Supp.2d 1 (D.D.C. 2006 ) (Despite the government’s opposition, the court appointed a special master, in light of the magnitude and complexity of the allegations, the volume of anticipated discovery, and parties’ inability to agree on basic scheduling matters.).

<sup>144</sup> *New Niche for E-Discovery: Special Master*, THE ESTRIN REPORT BLOG (Feb. 11, 2008, 5:10 PM), [http://estrinlegaled.typepad.com/my\\_weblog/2008/02/new-niche-for-e.html](http://estrinlegaled.typepad.com/my_weblog/2008/02/new-niche-for-e.html) (noting that a special master may charge \$500 per hour, and \$250 per hour for traveling expenses).

<sup>145</sup> *See* Petr Hajek, *Fuzzy Logic*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY, <http://plato.stanford.edu/archives/fall2010/entries/logic-fuzzy/> (last visited Apr. 13, 2012) (Fuzzy logic is a logic dealing with partial truth and approximation, instead of binary truth (true or false). Fuzzy logic is used in searching using linguistic terms.).

<sup>146</sup> Jeffrey I. Ehrlich, *Westlaw’s New Upgrade – Westlaw Next: Report of an Early Adopter*, WWW.PLAINTIFFMAGAZINE.COM (May 2010), [http://www.plaintiffmagazine.com/May10/Ehrlich\\_Westlaws-new-upgrade-WestlawNext\\_Plaintiff-magazine.pdf](http://www.plaintiffmagazine.com/May10/Ehrlich_Westlaws-new-upgrade-WestlawNext_Plaintiff-magazine.pdf).

## 1. Natural Language Search

Natural language search makes it easier to design search terms because the exact Boolean logic is not required in search term design: the user types the search terms using ordinary English.<sup>147</sup> The search engine, equipped with natural language search capability, will (hopefully) understand the search terms in the same way humans understand languages. Different approaches have been used to “understand” natural language, using artificial intelligence.<sup>148</sup> The retrieval result will be more relevant and the user can search in ordinary English, instead of using Boolean search terms or “query language.”<sup>149</sup> This approach minimizes difficulties in designing search terms, since the exact logic and phrases do not need be determined.<sup>150</sup>

Fuzzy logic behind natural language search provides for a “gray area,” where terms with similar meanings are included in the search. This helps to reduce some of the aforementioned problems with Boolean searches. One does not need to key in “strike” and then “striking” separately to return all emails containing these terms. Instead, “strike” alone would be enough in a natural language search, as the result will include “striking.” This “gray area” also allows for typing mistakes. Natural language searches can account for typing mistakes, thus producing relevant documents where keywords are misspelled.

While natural language search appears to be more user-friendly and give better result, the basis of its operation is still Boolean search logic.<sup>151</sup> In essence, fuzzy logic translates a phrase into complicated Boolean theorem.<sup>152</sup> For example, a natural language search for “all birds that live in Africa” is translated to something like (“bird\* + liv\* + Africa”).<sup>153</sup> Thus, natural language search still relies on elaborate keyword searches. Inevitably, the efficiency and accuracy of natural language search cannot escape the shadow of Boolean search terms.

---

<sup>147</sup> Edward A. Stohr et al., *Research in Natural Language Retrieval Systems 2* (Ctr. for Res. on Info. Sys. Computer Applications and Info. Sys. Area Graduate Sch. of Bus. Admin., N.Y.U. Working Paper No. IS-82-10, 1982).

<sup>148</sup> *Id.* at 4.

<sup>149</sup> *Id.*

<sup>150</sup> *See Gross*, 256 F.R.D. at 134 (The parties disputed search terms with Boolean logic, where the Boolean logic was essential to exclude irrelevant emails, while include relevant ones.).

<sup>151</sup> Baron, *supra* note 50, at 198.

<sup>152</sup> *Id.* at 202.

<sup>153</sup> *Id.*

As Boolean search method's efficacy is undetermined, natural language search efficacy is unlikely better than that of Boolean search.

## 2. Concept Search

Concept search is another search method attorneys have employed in an effort to achieve better results in retrieving ESI.<sup>154</sup> This search method aims at retrieving documents relating to the topic of the keyword without specifying the exact word.<sup>155</sup> A well-known example is documents about Eskimos and igloos are related to Alaska, even though they do not specifically mention the word "Alaska."<sup>156</sup>

Concept search is achieved by implementing a large set of rules, aiming at simulating semantic language.<sup>157</sup> In a database, rules govern the relationship between individual concepts, thus connecting different concepts.<sup>158</sup> Similar to natural language search, concept search is also based on abstract driven discovery rule, where the logic behind the rule is still binary.<sup>159</sup> In other words, concept search also translates the keywords into complicated Boolean logic and conducts the search based on a combination of rules. The efficiency of concept search has not been tested. At least one court has suggested concept search, as an alternative to keyword search.<sup>160</sup>

Despite attorneys' attempts, the efficiency of alternative search methods has not been tested. Due to the vast amount of information available in electronic form, the only way attorneys can sift through it all is to rely on search methods to retrieve documents. The risk of under producing always exists. While this problem also exists in tangible discovery, the risk of under production is higher in e-discovery. In sum, existing search methods place attorneys in a world where the bar in e-discovery is set too high. When diligent and competent attorneys risk sanctions due to technology, the bar that attorneys and parties are measured against in e-discovery must be reset. The proposed amendments to FRCP will reset this bar, reduce

---

<sup>154</sup> *Id.*

<sup>155</sup> *Id.*

<sup>156</sup> *Id.*

<sup>157</sup> Donald J. Berndt, *AX: Searching for Database Regularities Using Concept Networks*, 7 (Dept. of Info. Sys. Stern Sch. of Bus., N.Y.U. Working Paper No. IS-95-11, 1995) (semantic language is ordinary language).

<sup>158</sup> *Id.* at 2.

<sup>159</sup> *Id.*

<sup>160</sup> *See* Disability Rights Council of Greater Wash. v. Wash. Metro. Transit Auth., 242 F.R.D. 139, 148 (D.D.C. 2007).

discovery disputes, assure attorneys of discovery compliance, and ultimately reduce the cost of litigation.

## **IX. CONCLUSION**

Searching for responsive documents in data stored electronically is a challenge. Computers and electronics are common fixtures in everyday life. Changes in the way people store information result in changes to how attorneys sift relevant information from the irrelevant. In discovery, searching for information stored electronically can only be accomplished by implementing search methods. Search methods and keywords used are determinative of documents retrieved. Ultimately, the search method determines the produced documents.

With the amount of information available in electronic form and the importance of search methods in documents retrieval, FRCP amendments to address search methods are warranted. First, FRCP 26(f)(3)(C) should mandate parties' agreement on search methods in ESI searches. Second, FRCP 26(g)(1)(A) should allow for a certification process in which parties certify that they have searched using the agreed upon search method in specified databases.

As the information available in electronic form is of unprecedented magnitude, the burden of searching for responsive documents in ESI is different from discovery in the past. E-discovery also has unique characteristics, in which the search method is determinative of documents produced and the chance of an opposing party finding different documents is common place. Search methods are important in e-discovery even though the efficiency of keyword searches is low.

Given the challenges in searching for ESI, courts and attorneys have attempted different solutions, including appointing technical masters and using different search methods other than keyword search. These solutions do not address the underlying problem: the standard for complete production in discovery, as it stands right now, is infeasible in e-discovery.

With the proposed amendments, FRCP can address two difficult problems particular to e-discovery. It can force parties to consider e-discovery and its burden early in the suit to reduce the chance of discovery disputes, thus conserving judicial resources. Attorneys can feel more confident in signing their names to discovery replies, since they will be able to certify that the production is complete and stand by their certification without fear of the e-discovery monster.