

# Magistrate Judge Matthewman's New E- Discovery Paradigm and Solving the E-Discovery Paradox

*Professor William F. Hamilton*



## MAGISTRATE JUDGE MATTHEWMAN’S NEW E-DISCOVERY PARADIGM AND SOLVING THE E-DISCOVERY PARADOX

William F. Hamilton\*

*The exabytes of digital information streaming about us today are rich rivers of evidence that will help us find the truth and move us to do justice more swiftly, more economically and more honorably than ever before. It will require every litigator to master new skills and tools, and alter the approaches and attitudes we bring to the adversarial process. We must reinvent ourselves to master modern evidence or be content with a justice system that best serves the well-heeled and the corrupt. The path to justice is paved with competent evidence and trod by counsel competent in its use.*

- Craig Ball<sup>1</sup>

Magistrate Judge William Matthewman’s view from the bench is wary. Litigation is in the midst of a transition to electronically stored information with an undetermined outcome. The stakes are high. Litigation must produce a balance of correctness, fairness, and efficiency. At risk is legitimacy of the litigation itself.<sup>2</sup>

We are gradually leaving a quasi-dystopian era of “no-holds-barred” discovery slugfests featuring overreaching, recalcitrance, posturing, and exaggeration. On the future’s horizon, advance parties have made fitful, hesitant, and, at times, successful forays into the rich terrain of electronically stored information.<sup>3</sup> A brighter evidentiary world awaits, yet the courts and counsels seem stymied by obsolete practices that weigh “like a nightmare on the brains of the living.”<sup>4</sup> The more egregious e-discovery abuses have been corralled,<sup>5</sup> but our litigation practice seems stuck far below the pinnacle we can achieve. Judge Matthewman’s welcome article is a guide from the dark cave<sup>6</sup> of a previous era into the

---

\* Senior Legal Skills Professor, Levin College of Law, University of Florida.

1. Craig Ball, “*Whether I Shall Turn out to be the Hero of My Own Life, or Whether That Station Will be Held by Anybody Else, These Pages Must Show*,” BALL IN YOUR COURT (Aug. 20, 2011) <https://craigball.net/2011/08/20/hello-world/> [<https://perma.cc/R7RW-LY89>].

2. Jon O. Newman, *Rethinking Fairness: Perspectives on the Litigation Process*, 94 YALE L.J. 1643, 1646 (1985).

3. George Socha et al., *Cluster Clear: Are Clustering Tools the Solution to Tedious Identification and Reduction Processes in E-discovery?*, 101 JUDICATURE 14, 17 (2017); see, e.g., *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182, 193 (S.D.N.Y. 2012) (holding that computer-assisted review should be seriously considered in cases using large-data-volume sets).

4. KARL MARX, THE EIGHTEENTH BRUMAIRE OF LOUIS BONAPARTE 1 (1852).

5. William Matthewman, *Towards a New Paradigm for E-Discovery in Civil Litigation: A Judicial Perspective*, 71 FLA. L. REV. 1261, 1263–64 (2019).

6. E-Discovery failures related to preservation, collection, and production are becoming less dramatic and egregious. See, e.g., *Zubulake v. UBS Warburg LLC*, 220 F.R.D. 212, 222 (S.D.N.Y. 2003) (stating that UBS had a duty to preserve all of the backup tapes at issue but destroyed them with the requisite culpability). E-Discovery’s progeny has driven home

bright evidentiary light offered by electronically stored information. Judge Matthewman provides the foundational principles of a new electronic discovery paradigm to help us scale the final cliffs.

How did litigation get into this troubling discovery mess? The story starts with the revolutionary federal civil procedure rule changes in 1938. The 1938 amendments transitioned the federal courts from formalized pleading with minimal information exchange before trial to “notice pleading” where the broad-brush pleaded claims would be unpacked, explored, and refined before trial.<sup>7</sup> This new litigation discovery domain opened its cavernous jaws between the pleadings and trial with profound consequences. Suddenly “relevant” information *held by the opposition* became an available and valuable commodity. Ironically, however, the 1938 rule changes were designed to avoid trial surprise, ambush, and unfairness but nonetheless created a *discovery paradox* that has generated a tsunami of litigation. As a result of the 1938 amendments, a party may request information from the opposition through the various discovery vehicles, such as requests for production, interrogatories, and depositions, but compliance is dependent on the opposition finding and producing the requested information.<sup>8</sup> An information production obligation is imposed on the producing party, but the producing party obtains no “transactional” benefit from performance. The producing party must expend time and resources producing documents and responding to discovery requests but gains nothing from the disclosure of information in its possession.<sup>9</sup> Discovery becomes a coerced gift. Given this paradox, it is not surprising that production requests are often met with resistance, foot-dragging, narrow interpretations, questionable objections, intentional sluggishness, and outright non-compliance.<sup>10</sup> And,

---

counsel’s preservation duties. Craig Ball’s constant drumbeat for native production involving metadata is making native production more acceptable, Craig Ball, *A Guide to Forms of Production*, BALL IN YOUR COURT (May 19, 2014) <https://craigball.net/2014/05/19/a-guide-to-forms-of-production/> [<https://perma.cc/FQQ4-W6TK>], and production debacles such as producing only Tagged Image File Format (“TIFF”) images or flat PDFs or striping relevant attachments from emails is rarely seen. See *Abu Dhabi Commercial Bank v. Morgan Stanley & Co.*, No. 08 Civ. 7508 (SAS), 2011 WL 3738979, at \*2–3 (S.D.N.Y. Aug. 18, 2011). However, the remaining unbridled e-discovery gorilla in the room is search, analysis, and review. See [www.edrm.net](http://www.edrm.net) for additional e-discovery trends.

7. Charles E. Clark, *History, Systems and Functions of Pleading?*, 11 VA. L. REV. 517, 518–19 (1925); see also Rex R. Perschbacher & Debra Lyn Bassett, *The Revolution of 1938 and its Discontents*, 61 OKLA. L. REV. 275, 285 (2008).

8. See FED. R. CIV. P. 26–37.

9. I use the term “transactional” to emphasize the lack of “intrinsic” value to the producer of discovery performance. Of course, there are numerous potential externalities and collateral impacts that may come into play, such as the avoidance of sanctions, reducing costs, deterring the opposition, and encouraging cost savings through mutuality.

10. The structure resembles a contract that demands performance without any consideration. Such a “contract” creates no legally binding performance obligation; analogously, discovery obligations provide no transactional consideration for the producing party.

of course, whatever the information production, it is never sufficient for the gift recipient who believes that discovery is the gift that keeps on giving. Judge Matthewman refers to the constellation of these behaviors as “The Old Discovery Paradigm.”<sup>11</sup>

Discoverable information must be relevant,<sup>12</sup> but relevance is an elastic concept, and stretching and contracting the envelope of relevance became a litigator’s art form. Discovery can readily be resisted with objections and motions for protective orders. The vagueness and ambiguity of human language—as opposed to formal computer languages—lends itself to discovery objections. Virtually every discovery request can be portrayed by an overzealous litigator as ambiguous, designed merely to oppress, seeking irrelevant information, requiring the production of privileged information, overbroad, too expensive, too burdensome, etc. Human language is not a precise tool, and a savvy practitioner can transform the initial apparent clarity of any discovery request into a mishmash of confusion. Motions to compel invariably follow. Requesting parties fear that responding parties will engage covert and implicit tactics and respond with elaborate definitions, instructions, and requests which create yet another escalating cycle of resistance. Too many litigators are driven by the premise that the opposition holds critical information hidden behind objections, clever attorney phraseology, and an exceedingly narrow reading of discovery requests. Judge Matthewman also refers to this discovery chaos as “The Old Discovery Paradigm.”<sup>13</sup>

Discovery battles are fought because the stakes are high. After 1938, litigation increasingly became a struggle over information exchange. Information provided the substance and shape of issues for trial and the proofs that would be offered at trial. A determinant of “winning” became controlling the flow of information: blocking access to your information and gaining access to the opposition’s information. A dysfunctional dynamic emerged where each side sought to maximize “one-way” discovery. As reflected in Judge Matthewman’s article, judicial officials dreaded wading into this ill-defined morass and often expressed exasperation, frustration, and disappointment.<sup>14</sup> Discovery sanctions became a jurisprudence of judicial anger.<sup>15</sup>

As Judge Matthewman nicely articulates, paper-era baked-in litigation impulses and extreme conduct wrecked even greater havoc when information became predominantly electronically stored information.<sup>16</sup>

---

11. See Matthewman, *supra* note 5, at 1263–65.

12. FED. R. CIV. P. 26(b).

13. See Matthewman, *supra* note 5, at 1263–65.

14. *Id.* at 1264–65.

15. *Id.* (“Judges . . . can no longer tolerate such discovery shenanigans.”).

16. *Id.* at 1265.

While the 1938 rules change shifted the operational locus of civil litigation from the trial to written discovery and depositions, the emergence of electronically stored information moved the operational locus of civil litigation from written and deposition discovery to “document” exchange, i.e., requests to produce increasingly valuable electronically stored information. The fulcrum of litigation was no longer trials, nor even deposition testimony, but electronically stored information that was often obtained at the earliest stage of litigation.<sup>17</sup>

Document requests were, of course, important in late 20th century litigation, but the volume of produced documents was typically meager, and the documents themselves often unrevealing. For example, document requests frequently did not unearth potentially valuable discarded paper drafts. Left in the file cabinet was only the refined, sanitized, and polished final product. The actual story behind the document had to be reconstructed from numerous lengthy and tedious depositions which linked together memory fragments from deponents typically not inclined to recall unpleasant or unfavorable events.

The meteoric rise of electronically stored information in the 21st century changed all this. Suddenly, smart litigators could build their cases utilizing emails, notes, drafts with tracked changes, social media posts, images, videos, and text messages that are created virtually contemporaneously with events. This evidentiary windfall however came with a price: The relevant information needed to be extracted from large volumes of mostly irrelevant information utilizing digital information retrieval techniques. Litigators with virtually no experience in computer science or information retrieval or law school training in electronic discovery found themselves in an alien world of bits, bytes, gigabytes, megabytes, and terabytes of data.<sup>18</sup>

As Judge Matthewman notes, many e-discovery failures arise from litigators carrying their venerable paper-based discovery habits into the new e-discovery arena with disastrous results.<sup>19</sup> Not only was valuable electronically stored information missed, lost, or destroyed, but worse, attorneys applied the old mindset of discovery resistance. Some attorneys, eager to resist discovery and increase the opposition’s cost, objected to forms of production that contained valuable information, intentionally used wrong-headed and untested search terms, obliterated metadata, insisted on manual reviews of terabytes of data, and failed to

---

17. Early case investigations no longer required reliance on a witness’s biased and subjective memory. The witness’s emails, text messages, word processing documents, social media posts, and cell phone data tell a clearer and probably more reliable story.

18. A large case in the 1980s might consist of 50 Bankers Boxes of paper. E-Discovery volume in a medium to small case in 2010 might easily consist of 50 gigabytes of data, or roughly 1,500 Bankers Boxes!

19. See Matthewman, *supra* note 5, at 1272.

preserve relevant and important information. For example, one such tactic was to produce the requested information to the opposition in one unsearchable PDF file consisting of thousands of files converted from their native format.<sup>20</sup>

Rather than recognizing electronically stored information as a treasure trove of available information, litigators in the early days of the e-discovery era<sup>21</sup> continued their traditional focus on depositions and often used document requests simply as a cudgel to drive up litigation costs and burdens. The view of electronic discovery as largely an unnecessary burden is reflected in the 2006 and 2016 e-discovery changes to the Federal Rules of Civil Procedure. These rule changes have often been interpreted to impose a priori limits on so-called out-of-control e-discovery: no discovery of not reasonably accessible information, no discovery of duplicative information, no disproportionate discovery.<sup>22</sup> Some courts have tried to “control” and “restrain” e-discovery with rules and orders limiting the number of search terms or providing mandatory document production stages. However well-intentioned, such arbitrary shackles miss the point: Electronic discovery will only advance with increased competence. The problem is not that we have too much information: The problem is that attorneys do not understand electronically stored information and are not adept at using available information retrieval and analytical technologies to find case-important documents.

The modestly increasing electronic discovery competence that Judge Matthewman has experienced over the past ten years is driven by three principal sources with varying impacts. The first is sanctions. Sanctions are available to the court under Federal Rules of Civil Procedure 16, 23, and 37. Unfortunately, rule enforcement has its place, but it will never alone be sufficient. Sanction motions face significant hurdles and fear of sanctions are only one factor governing litigation behavior. Moreover, fear of sanctions only produces minimal competency at best, doing just enough to withstand modest scrutiny. Another driver of e-discovery competence is the simple fact that litigators who know electronic discovery can better advance their clients’ litigation goals. E-discovery competence wins cases; however, the correlation between competence and success is not always readily apparent, especially at the commence of a case. Moreover, less than competent e-discovery attorneys are oblivious to the impact of their limitations.

Finally, the third driver of e-discovery competence is the litigator’s

---

20. *In re Seroquel Prods. Liab. Litig.*, 244 F.R.D. 650, 665 (M.D. Fla. 2007).

21. As a discipline, e-discovery might broadly be considered coincident with the 21st century, notwithstanding that important cases in the late 20th century involved data exchange and analysis. The Sedona Conference’s first e-discovery publications were in 2000.

22. *See* FED. R. CIV. P. 26(b)(2)(C).

clients. From the client perspective, litigation is most often a question of risk management involving time and money. The standard American rule requires the parties front their own discovery expenses.<sup>23</sup> Too often clients perceive resistance to discovery as beneficial in the short run as money which has often not been budgeted for litigation is saved. Assuming sanctions are avoided, it is often cheaper to pay the relatively modest attorney's fees incurred in resisting discovery than to pay the costs of document retrieval, review, and production, or for deposition preparation and defense. Moreover, the case may settle, or the opposition may wither. Money not spent is money saved.

The good news—from the perspective of e-discovery competence—is that client perceptions of the discovery risk calculus are slowly changing in the e-discovery era. A lost discovery battle in the paper era would not have likely triggered substantial costs rivaling the amount in dispute. In the e-discovery era, repairing preservation failures, collection failures, search failures, and production format failures can be very expensive. Moreover, clients are quickly learning that e-discovery costs have less to do with the volume of data than with counsel's knowledgeable deployment of advanced technologies. E-discovery mistakes and “do-overs” are cost black holes that clients and in-house counsel have learned to dread. However, client pressure to minimize cost is not a constraint that produces genuine e-discovery competence.

Thus, the time is ripe for a wholesale adoption of Judge Matthewman's new E-Discovery Paradigm. The New Paradigm is composed of the following triad of “Core Components.”

#### A. E-Discovery Competence and Capabilities

- #1. Proper and timely preservation.<sup>24</sup>
- #2. Prompt and complete Rule 26(a)(1) disclosures.<sup>25</sup>
- #3. Targeted and precise discovery requests.<sup>26</sup>
- #7. Early and routine involvement of the parties' in-house information technology (IT) professionals during the discovery process.<sup>27</sup>
- #8. Use of e-discovery companies, vendors, and experts to assist as needed in litigation.<sup>28</sup>
- #9. Greater reliance on technology assisted review (TAR), search strings, sampling, artificial intelligence, and other scientific or technical methods to aid, hasten, and economize the discovery process in a

---

23. See BRITTANY K.T. KAUFFMAN, ALLOCATING THE COSTS OF DISCOVERY: LESSONS LEARNED AT HOME AND ABROAD, INST. FOR ADVANCEMENT AM. LEGAL SYS. 22 (2014).

24. Matthewman, *supra* note 5, at 1266.

25. *Id.* at 1268.

26. *Id.* at 1269.

27. *Id.* at 1274.

28. *Id.* at 1275.

transparent and reliable manner.<sup>29</sup>

### B. E-Discovery Reciprocity and Fair Dealing

#4. Complete discovery responses devoid of boilerplate and meritless objections.<sup>30</sup>

#5 Professionalism, cooperation, and honest good faith, personal conferral among the parties' in-house counsel when the inevitable discovery dispute arises.<sup>31</sup>

### C. E-Discovery Management and Engagement

#6. Limitation of discovery by the court to discovery that is relevant and proportional to pending claims of defenses as required by Rule 26(b)(1), and elimination of wasteful or unnecessary discovery.<sup>32</sup>

#10. Active participation of judges in the discovery process and prompt resolution of any discovery disputes by the court.<sup>33</sup>

Judge Matthewman's New Paradigm is not a mere tinkering with the existing rules. Do not under-estimate. The New Paradigm mandates sweeping and fundamental changes in discovery practices. Judge Matthewman offers us a new litigator's mind-set, a new litigation weltanschauung, a new totalizing gestalt. Perhaps the depth of the change can be captured by recalling that the original expression "*paradigm shift*" was coined by Thomas Kuhn in his landmark book *The Structure of Scientific Revolutions*.<sup>34</sup> Kuhn theorized that the advance of scientific knowledge is not merely a knowledge accretion build on a solid permanent conceptual framework.<sup>35</sup> Kuhn demonstrated that science advances in leaps that challenge pre-foundational concepts.<sup>36</sup> For example, after Copernicus, everything we knew about the cosmos changed.<sup>37</sup> Every new scientific world view represents a group or constellation of fundamental concepts that give new coherence and understanding to our experience. Modern science is a tale of such paradigm shifts.

Judge Matthewman's New E-Discovery Paradigm articulates a similar fundamental paradigm shift in litigation whose Core Components are the interrelated and interlinking fundamental concepts. For example, Core

---

29. *Id.* at 1276.

30. *Id.* at 1270.

31. *Id.* at 1271.

32. *Id.* at 1276.

33. *Id.* at 1279.

34. THOMAS S. KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* 85 (2d ed. 1970).

35. *Id.* at 3.

36. *Id.*

37. *Id.* at 74.



Component #1, preservation, is required by Core Component #2, Rule 26, which likely requires activation of Core Component #4, involvement of IT professionals. None of the Core Components stand alone as a solitary beacon. Rather, each Core Component must be understood as a foundational element of the new Matthewman paradigm. The ten Core Components are one interlaced mosaic, all of which together comprise the new paradigm shift revolving around the interwoven triads of competence, reciprocity, and engagement of counsel, courts, and parties.

What does this new paradigm shift have to say about our pesky *discovery paradox that wrought so much havoc in the paper era*? Let's see how the new paradigm resolves the *discovery paradox* by unpacking the implications of Judge Matthewman's Core Component #9.

*Greater Reliance on Technology Assisted Review (TAR), Search Terms, Sampling, Artificial Intelligence, and Other Scientific or Technical Methods to Facilitate the Discovery Process in a Transparent and Reliable Manner.*<sup>38</sup>

Keep in mind the origin of the *discovery paradox*. The responding party has no transactional incentive in information production. The responding party is already in possession of the information. The information production does not generate an information advantage. Rather, it creates a one-way information transfer often causing discovery malfeasance. Does competently practiced electronic discovery under the Matthewman New E-Discovery Paradigm offer any solution to the *discovery paradox*? Surprisingly, the answer is "yes."

E-discovery is a classic information retrieval binary classification problem.<sup>39</sup> The document (electronic file) is responsive to a production request or not. Two groups are created: documents to be produced subject to privilege review and documents to be withheld as not responsive.<sup>40</sup> The producing party is highly sensitive to what is being produced. Quality control must ferret out privileged documents and the producing party must be keenly aware of key data being produced. Unfortunately, the unproduced pile gets less attention. Judge Matthewman correctly focuses on TAR, machine learning that predicts and ranks likely responsive documents. Keyword search technology,<sup>41</sup> however, also "predicts"

---

38. Matthewman, *supra* note 5, at 1276.

39. DOUGLAS W. OARD & WILLIAM WEBBER, INFORMATION RETRIEVAL FOR E-DISCOVERY 137 (2013).

40. *Id.*

41. Keyword search is just as much a predictive tool as more sophisticated supervised and unsupervised machine learning. The presence of a keyword only predicts that the document may

responsive documents as well.<sup>42</sup> Surprisingly simple measurement tools are now available to the requesting party *to manage* the producing party without any invasion of attorney work-product. This is the new jurisprudence of search validation. The requesting party's principal concern is typically recall: Has the opposition identified and produced the responsive documents by whatever search and analytic technology is employed, regardless of whether the technology is keyword search strings, TAR, or other technology? Too many litigators responding to the reasonable question of whether their search is complete dismissively retort, "Of course our production is good, we used the right search terms or TAR process or other technology!" However, the question of whether the production is complete is not an attack on counsel's integrity, but simply asks whether the search results were tested by looking not at what was produced but at what was left behind. Were the documents and files that were determined to be non-responsive tested to gauge whether any scope of non-responsive documents "eluded" detection? Was an *elusion* test performed? A search is only as good as what was not left behind.

So, let us follow the Matthewman New E-Discovery Paradigm and require sampling of the *non-produced* documents.<sup>43</sup> While the nuances of sampling are beyond this response, the basic idea is clear: a review of a very small random sample will provide a high level of confidence even for extremely large collections of electronically stored information. . Producing parties should be required to certify elusion testing. Sampling is quick, inexpensive, and can provide the *requesting party* with a reasonable level of confidence in the overall validity of the search methodology.

## CONCLUSION

Judge Matthewman's New Paradigm for E-Discovery is a mandatory

---

be responsive. No search in the field obtains 100% precision: false positives abound. A so-called false-positive is just another way of saying the prediction was not accurate.

42. CHRISTOPHER D. MANNING ET AL., AN INTRODUCTION TO INFORMATION RETRIEVAL 118–19 (2009). Numerous bases are available to predict and rank documents. For example, Tf-idf, term frequency-inverse document frequency, predicts the importance of a document based on keyword frequency within a document and the frequency of documents in the collection containing the keyword. *Id.*

43. See *City of Rockford v. Mallinckrodt ARD Inc.*, 326 F.R.D. 489, 496 (N.D. Ill. 2018); see also *In re Mercedes-Benz Emissions Litig.*, No. 2:16-cv-881, 2019 U.S. Dist. LEXIS 193948, \*7 (D.N.J. Nov. 4, 2019). The goal is not to establish some numerical standard, but to require testing of the null set and certification that testing identified no responsive documents. If testing does identify responsive documents, then either the search can be refined or the producing party can produce and argue that the additional located responsive documents in the test are of marginal value and thus the search is successful.

prescription. Among its many immediate benefits is the solution of the discovery paradox. The electronically stored information toolbox now allows the requesting party to manage the performance of the producing party, thereby solving the discovery paradox that has frustrated lawyers and judges since the 1938 inception of broad document discovery. Courts and lawyers should insist upon elusion testing and banish the discovery production obstructions induced by the discovery paradox to the proverbial dustbin of history.