Using Technology to Facilitate Production of E-discovery

Timothy J. Pramas
Allison Berres
Shannon L. Bjorklund
Steven J. Ellison

Follow this and additional works at: http://open.mitchellhamline.edu/wmlr

Recommended Citation
Available at: http://open.mitchellhamline.edu/wmlr/vol40/iss2/8
USING TECHNOLOGY TO FACILITATE PRODUCTION OF E-DISCOVERY

Minnesota E-Discovery Working Group 5†

INTRODUCTION .................................................................................. 591
I. EARLY DATA ASSESSMENT .......................................................... 592
   A. Introduction to Early Data Assessment ..................................... 592
   B. Technology Functionalities and Options for EDA ................. 595
   C. Strategic Considerations for Choosing an EDA Approach and Choosing Technology in Specific Cases .......... 597
   D. Legal Obligations: How Much Must You Disclose to Opposing Counsel About Your EDA Approach (and Do You Want to Disclose Even If It Is Not Required)? .......... 600
   E. Additional Sources of Information Relating to Early Data Assessment ................................................................. 607
II. EFFICIENT AND DEFENSIBLE REVIEW ....................................... 608
   A. Cost of Review ......................................................................... 608
   B. Managed Review ..................................................................... 609
   C. Roles and Responsibilities .................................................... 622
III. PRODUCTION FORMAT .............................................................. 625
   A. Production Specifications ...................................................... 625
   B. Standard Production Requests ............................................. 625
   C. Production of Specific File Types .......................................... 625
   D. Final Thoughts About Production ......................................... 627
IV. SOCIAL MEDIA AND E-DISCOVERY ............................................. 628
   A. Social Media and ESI ............................................................. 628
   B. Potential Social Media Evidence Uses ................................... 629

† The Minnesota E-Discovery Working Group is a grassroots organization that was founded in 2011 with the goal of writing five separate papers that address various aspects of e-discovery best practices from a Minnesota perspective and could be used as a resource by both judges and lawyers in Minnesota. Its members consist of members of the Minnesota judiciary, in-house attorneys, attorneys practicing with law firms across Minnesota, and e-discovery experts. The Working Group and the William Mitchell Law Review thank Briggs & Morgan, P.A. and Fredrikson & Byron, P.A. for their financial contribution to this joint project.
C. Discoverability of Social Media Posts ........................................... 630
D. Requests for Production ............................................................ 631
E. Recovery of Social Media Evidence ............................................. 633
F. Computer Hard Drive Data Recovery ......................................... 634
G. Social Media Visualizers ............................................................. 634

V. SMART PHONES AND TABLETS: ESI FOR REVIEW ............... 635
A. Types of Content on Portable Devices ........................................ 635
B. Other Content—Metadata ........................................................... 636
C. Another Data Source—Service Provider Business Records ....... 636
D. Another Data Source—Call Detail Records (CDR) ................. 638
E. Text Messages ........................................................................... 639

VI. CONCLUSION ........................................................................... 640
APPENDIX A: PRODUCTION FORMATS ....................................... 641
APPENDIX B: HOW IS E-DISCOVERY DIFFERENT FROM
DIGITAL FORENSICS? ................................................................. 643
APPENDIX C: SUBPOENA TO FACEBOOK ................................... 644

About the Authors††

Timothy J. Pramas (Co-Chair) is Senior Associate General
Counsel for the University of Minnesota’s Office of General
Counsel. Tim previously worked as an attorney at Manty &
Associates, P.A., where he represented clients in a wide variety of
civil litigation and was a frequent lecturer on e-discovery issues.
Tim also worked at SUPERVALU INC. as a Litigation Attorney
where he spearheaded many of the company’s electronic discovery
initiatives.

Allison Berres is the Director of Client Solutions for the Client
Technology Group at Faegre Baker Daniels. Allison manages the

†† The views expressed herein do not necessarily reflect the views of the
firms and companies listed or the Minnesota State Judicial Branch. Working
Group 5 wishes to acknowledge the contributions of John J. Carney, who
contributed content for the social media and mobile technology portions of this
paper. Working Group 5 also wishes to acknowledge co-chair Caroline Bourdeau
Sweeney, of Dorsey & Whitney LLP, and committee members Michael Dewberry,
of Blue Cross Blue Shield of Minnesota; Mary Harens, attorney at Mary Harens
Law Office; Kristin Lehmann, attorney at Knoll Onrack, Inc.; Michael J. McGuire,
attorney at Littler Mendelson; Matthew Pederson, of Dell Compellent; and Patrick
Williams, attorney at Briggs & Morgan, P.A., for their contributions to the
Working Group.
delivery of litigation support and e-discovery services. She is responsible for developing and managing “best practices” protocols and policies in e-discovery legal practice, process, and technology.

Shannon L. Bjorklund is an attorney at Dorsey & Whitney L.L.P.. She focuses primarily on litigation involving the healthcare industry, including intellectual property litigation, and counseling public and quasi-public entities on document retention, discovery obligations, and public records law.

Emily E. Duke (Co-Founder of the Minnesota E-Discovery Working Group) is a shareholder at the international law firm of Fredrikson & Byron, and co-chairs the firm’s Electronic Discovery Resources Group and its Data Protection and Cybersecurity Group. She advises clients on data security, mitigating risk, and protecting business know-how, trade secrets, and assets.

Steven J. Ellison the owner of Ellison Law Offices, P.L.L.C., has served on the Steering Committee for the Defense Research Institute’s section on product liability and e-discovery issues. He routinely counsels clients on electronic discovery, data management, and data-storage issues and is a frequent speaker on those subjects.

Andrew Hansell is an Associate Vice President in the Legal Services Outsourcing Division of CPA Global. He works with corporate clients to manage end-to-end document collection and review in government investigations, business transactions, and civil litigation, including as a member of the national counsel team for a manufacturer involved in an industry-wide product liability multidistrict litigation.

Sean Harrington is a digital forensics consultant for his local firm, Attorney Client Privilege, L.L.C. and also a cybersecurity policy analyst and risk assessor for US Bank. Sean has served on the board of the Minnesota Chapter of the High Technology Crime Investigation Association, is a current member of Infragard, the Financial Services Roundtable legislative and regulatory working groups, FS-ISAC, and is a council member of the MSBA’s Computer & Technology Law Section.

David S. Miller is a Senior Discovery Consultant at Falcon Discovery, Inc. In that position, he advises and assists Falcon’s clients on all aspects of the e-discovery process, from litigation holds, collections, searching, reviewing, and production.

Frank L. Nelson is the E-Discovery Manager at the law firm Nilan Johnson Lewis, PA. Frank has served on the board of the
Minnesota Association of Litigation Support Professionals and is a regular speaker at industry conferences and association meetings covering topics related to e-discovery.

Kimberlee Nelson has more than twenty years of experience managing complex litigation matters and developing methodologies and best practices for all stages of the e-discovery reference model (EDRM). In her current role as an independent e-discovery consultant, she partners with her clients to strategically manage current matters and proactively prepare for future litigation.

Patti Ray is Vice President at Carney Forensics where she manages sales and marketing, operations, and customer service. She works with corporate clients, legal firms, and top private investigation agencies nationwide to build partnerships that deliver mobile, computer, and online evidence strategies.

Sean Reisman is a managing attorney at CPA Global, specializing in e-discovery project management, litigation drafting, and legal research. In addition to managing large-scale e-discovery matters, he advises clients on efficient work flows and review technology.

Jennifer G. Swanton is Discovery Counsel for Medtronic, Inc. where she works on discovery issues arising from various legal matters including IP, product liability, class actions, and shareholders suits. Jennifer has extensive experience in the e-discovery vendor space and currently manages an e-discovery certificate program for the Organization of Legal Professionals.

Megan A. Thwaites is Litigation Counsel for the Chief Litigation Office at Boston Scientific Corporation. Meagan manages the Litigation Support team, which handles litigation and discovery efforts and is also responsible for the management of Boston Scientific’s electronic discovery practices, data processing and document review vendor relationships, and in-house electronic discovery tools.

INTRODUCTION

Group 5 of the Minnesota E-Discovery Working Group focused on technology to assess, review, and produce data. Chronologically this phase occurs after data are preserved and collected, which are topics addressed by other groups. This paper is divided into five parts and discusses the ways in which technology can be used to facilitate discovery of information that may be stored on electronic
devices. More specifically, the five sections address (1) early data assessment, (2) efficient and defensible review of electronic data, (3) production of electronic discovery, (4) issues pertaining to review and production of social media and electronic discovery, and (5) the fast-evolving challenges posed by smart phones, tablets, and the electronically stored information (ESI) found on these devices.

I. EARLY DATA ASSESSMENT

A. Introduction to Early Data Assessment

1. What Is EDA?

Early Data Assessment (EDA) means different things to different people. For those in the technical area of electronic discovery, it can mean sampling data; to others, it can mean a quick pass over the data in order to see what is there (i.e., a first review). Regardless of who is defining EDA, everyone agrees it can be a powerful tool for narrowing the scope of data to be preserved, reviewed, and ultimately produced in litigation.

After collection of relevant preserved data, EDA helps attorneys narrow the scope of data to be processed or reviewed. It looks at a large set of data early on, so as to help attorneys focus their processing and review of that data. EDA helps attorneys:

- Triage data by level of importance or relevance.
- Gain early visibility into the data collected, before processing or review.
- Improve efficiency for reviewing (or not reviewing) massive quantities of duplicate or near-duplicate documents.

EDA is different from Early Case Assessment (ECA). ECA involves analysis of the entire case, including the case merits versus cost effectiveness, not just documents and data that will potentially be the subject of discovery. ECA can encompass fact finding, venue analysis, damages assessment, liability analysis, investigation of opposing parties and counsel, litigation budget forecasting, and more.

2. Why Should I Use EDA?

EDA has many benefits:

- **Collection, review, and production:** EDA allows for more efficient collection, review, and results in a production. It allows parties to gather more of the responsive material, while reducing the cost of the production by eliminating irrelevant, duplicative, or unnecessary information. EDA assists with:
  - Providing robust reporting and allowing attorneys to slice and dice the data many ways to see patterns.
  - Allowing attorneys to triage the order in which to review data, starting with the most important data or custodians first.
  - Identifying potential custodians or identifying individuals who should not be part of the collection.
  - Determining whether additional collections are appropriate.
  - Gaining a better understanding of key case information (e.g., key ideas, e-mail threads, chains of communication, and connections between custodians).
  - Performing keyword analysis and refinement, to ensure a more efficient and effective collection, review, and production.

- **Communications and conferences with opposing counsel:** EDA allows an attorney to better communicate with opposing counsel and to adopt beneficial positions that will save the client’s time and money. EDA assists with:
  - Facilitating the required discovery conference with opposing counsel pursuant to Federal Rule of Civil Procedure 26(f) and Minnesota Rule of Civil Procedure 26.06, by allowing an attorney to better understand the format and scope of the data.
  - Selecting agreed-upon search terms that will not result in excessive numbers of false hits.
  - Selecting agreed-upon search terms that will adequately identify responsive documents.
  - Defending the search protocol and the party’s positions with regard to discovery stipulations and productions.

- **Use of documents for later litigation tasks (e.g., depositions, trial):** EDA allows attorneys to more efficiently and effectively identify
key documents that will play a major role in depositions, dispositive motions, and trial. EDA assists with:

- Identifying words and concepts important to the case in order for creation of effective keyword searches. This can be useful for document production and also for later searches and document organization, such as deposition preparation, summary judgment briefing, and trial preparation.
- Reducing review costs by eliminating irrelevant documents: Effective searches provide more targeted review, which reduces the scope of the review (by eliminating irrelevant documents) and increases the effectiveness of the review (by successfully identifying important documents).
- Allowing multiple views of the data, depending on the purpose of the review: For example, some tools allow for one platform with similar interfaces for preview and full review.

3. When Should I Use EDA?

EDA can be useful in many different contexts. Legal teams often use EDA for litigation, but it can be a powerful tool in many other types of matters. For example:

- Regulatory matters: To provide a quick overview of the scope of the matter, the types of documents the client possesses, and the key individuals who participated in communications about the subject matter.
- Litigation matters: To reduce data volume prior to document review, make document review more efficient (by triage or grouping), assist counsel in selecting meaningful agreed-upon search terms, and organize documents for later document-heavy tasks such as deposition preparation, summary judgment briefing, and trial preparation.
- Internal investigations: To take an early look at data and determine the key players, timing, and patterns of correspondence.
- Policy audits: To develop a “Tickler System,” and to understand compliance with internal policies and potential risk.

2. A tickler system can take many forms. For example, a tickler system may
B. Technology Functionalities and Options for EDA

There are an increasing number of options for products and technologies. Different technology is appropriate for different cases, as will be described in further detail in Part III. The legal team should consider the size of the case (both in monetary risk/reward and size of the data set) as well as time constraints and other factors when selecting the appropriate technology. This section provides an overview of some key functionalities that can be enormously helpful and cost-saving in the right case.

1. Search and Filtering Technology

Searching and filtering tools can be huge time-savers in high-volume cases. Below are descriptions of different searching and filtering technologies that legal teams should consider.

- **Traditional searching**: Traditional searching typically uses terms and connectors, proximity, wildcard, and expander searching. These tools can be very useful in EDA. Search queries can be run on multiple fields and can be done by keyword, use, various date options, metadata, and coding fields. These will allow the user to estimate the number of documents that will be generated using different combinations of key words. Stemming, phonic, synonym, related, fuzzy, multiword, and Boolean searches may also be used. Data and evidence bookmarking support categorization and organization.

- **Topic grouping and concept searching**: This type of searching uses intelligent technology and complex algorithms to group similar documents by concept or topic. This will allow the reviewer to quickly and accurately evaluate volumes of documents for relevance and responsiveness. This is also a very useful tool when reviewing productions and is a great alternative to traditional keyword searching.

- **Filtering**: Filtering can be done by type of document. For example, a filter may identify certain file types, encrypted files, decrypted files, duplicate files, near-duplicate files, 

provide prompts based on certain time frames (e.g., two weeks prior to a financial reporting deadline). Alternatively, a tickler system may be used to identify at-risk communications, such as communications with certain keywords or between particular groups or individuals.
e-mail chains, graphics, sender, recipient, subject line, date range, and source.

2. De-Duplication

Many document collections involve huge numbers of duplicate documents. According to some estimates, de-duplication reduces the size of a document collection by approximately fifty percent.³

There is more than one kind of de-duplication. For example, global de-duplication means that the technology will compare documents collected from all custodians and will retain only one copy of the document.⁴ By contrast, custodian de-duplication will retain one copy for each custodian.⁵

- **Near-duplicate document comparison:** This technology allows users to compare differences of extracted text, annotations, categories, or reviewer comments between two near-duplicates.

- **Duplicate, near-duplicate, and e-mail chains:** Different tools can remove duplicates, and group near-duplicates and e-mail chains together. De-duplication can be done globally or by custodian, and it can split documents or keep them in families.

3. Generating Reports and Other Analytic Tools

Reports and analytic tools can be enormously helpful. An attorney should carefully consider the features and advantages of each type of program before selecting a program for a specific case. Below are a few examples of analytic tools that may be available:

- **Reports:** Users can generate reports that provide information about many different categories. Examples include: overall data set size; number of sources; types of files (e-mails, .doc files, .pdf files, Excel spreadsheets, etc.); data profile; peaks and valleys in volumes of e-mail traffic; common e-mail subjects; and e-mail traffic between specified custodians.

- **Search and filter logs:** Users can determine whether the software retains detailed logs of searches or filters run and the results

---


⁵. Id.
thereof. These historical logs may be useful when a user wants to re-create the results of a previous search. Users should also determine whether the program can generate a user-friendly report of the searches/filters and the results thereof.

- **E-mail thread analysis and analytics**: Analytics provide users with the power to analyze and interact with their most potentially relevant data by visually displaying who is communicating with whom, when they communicated, and about which topics they were communicating. Early detection of common themes within the data set can be revealed, ultimately allowing document review teams to be more productive and accurate in making decisions regarding responsiveness and privilege.

- **Data dictionary**: A comprehensive data dictionary of all words in a data set can be a useful tool. This can enable the user to determine which key terms are important in the review, and provides the option to use this information in negotiations regarding search terms.

### C. Strategic Considerations for Choosing an EDA Approach and Choosing Technology in Specific Cases

EDA can be helpful in nearly every case, but the appropriate cost-effective tools will vary from case to case. This section provides a list of considerations when picking the appropriate technology for a given case.

#### 1. Choosing an Approach

- **Assessment of e-discovery experience and sophistication of case participants**: Everyone comes to e-discovery from a different place. This includes the client, the court, opposing counsel, and the legal team performing the collection and review.

- **Choosing who will manage the data assessment**: E-discovery—like all discovery—requires a balance between proportionality/cost considerations and a reasonably comprehensive, defensible effort to locate and produce documents. When choosing the party to collect documents, these cost and competency concerns are particularly important.
  - **In-house**: Some clients prefer to do data collection in-house. The major perceived benefit is cost reduction. Some companies have extremely sophisticated legal and technical teams, and other companies have virtually no
resources. Things to consider when deciding whether the collection should be done in-house include:
- Management;
- Overhead/burden;
- Technology changes during critical time period;
- Infrastructure and document retention policy (e.g., servers, backup tapes, shared drives);
- IT/legal team experience and ability;
- Repeatability/frequency of need.

- **Outside counsel**: Outside counsel may be able to collect and process the documents more efficiently. The major perceived benefit to this approach is to have the individuals who will be guiding and processing the case also oversee the document collection and assessment. The factors to consider for outside counsel are the same as for an in-house collection:
  - Management;
  - Overhead/burden;
  - Technology changes during critical time period;
  - Infrastructure and document retention policy (e.g., servers, backup tapes, shared drives);
  - IT/legal team experience and ability;
  - Repeatability/frequency of need.

- **Vendor**: The major perceived benefit to this approach is having an entity that specializes in data collection and management perform the collection and assessment of the data. Things to consider when deciding whether a vendor should do the collection include:
  - Management;
  - Cost;
  - Vendor’s experience with similar cases;
  - “Reinventing the wheel”,

---

6. Either a party or a vendor may have inefficient processes that result in unnecessary duplication—“reinventing the wheel.” For example, using a vendor may eliminate duplication because the vendor has established systems already in place that the client would otherwise need to create. In this instance the client would be “reinventing the wheel,” and it would be more efficient to use a vendor.
Prior relationship and past history with vendor.

- **Assessment of scope.** What does the case need, and what will the case support? Using an expensive EDA tool is not appropriate for all cases. Sometimes all that is needed is a quick look at the data. The team may need to know only the number of documents to be reviewed or the communication networks. This information can be gathered by looking at a few sample e-mails. There are also other ways you may perform EDA when your case does not support an expensive tool. For example, after interviewing custodians of data, you might learn that certain custodians file their electronic documents by project, in which case you may simply gather the documents that the custodian already segregated for the project at issue. Or, you might decide to run some basic searches or sample certain custodians, based upon information gleaned from interviews, to see where you are most likely to find the most relevant documents.

- **Consideration of timetable.** Do you have time to use an EDA tool? Depending on the turnaround time you have, sometimes using an EDA tool is simply not possible. For example, if your client must respond to a government subpoena within an extremely tight window, you may not have time to load the data onto the EDA tool, analyze it, and then review it. If time permits, using an EDA tool should start during or after data collection and end as the document review process begins in earnest.

### 2. Choosing the “Right” Technology

- **Assessment of data:** The volume and type of data (e.g., e-mails, engineering plans, audio files, etc.) greatly influence the tool that one should use. The tool analysis needs to include an assessment of the tool’s capabilities. For example:
  - Can the tool handle the data types in your data set?
  - Can the vendor process and provide access to the data set in the timeframe you need?

However, if the client has sophisticated retention, storage, or collection capabilities, a vendor may be “reinventing the wheel” when it attempts to collect, categorize, and search across the client’s data.
• **Cost/benefit of performing EDA:** How much processing time/expense will be saved if you can cull the data collected beforehand? How much will the EDA tool cost, and how much time will it save you in review? This analysis needs to include a look into how the data were collected and if you anticipate gaining enough information or reducing your data set enough to warrant the cost. One example that is often seen where EDA does not provide enough of a benefit is when a company is able to do some filtering in-house. When a company is able to pre-cull its data, you may not see enough of a benefit to move forward with EDA.

• **Functionality versus price of EDA options:** Does extra functionality make up for increased price? In some cases—particularly in cases involving a small number of documents or a small amount of money—the additional functionality may not be cost effective. Extra functionality will be particularly useful for cases involving large amounts of documents. The additional cost for the functionality may not be advisable in cases involving few documents or in cases where the parties are anxious to keep costs low (due perhaps to a smaller amount of money at stake).

D. *Legal Obligations: How Much Must You Disclose to Opposing Counsel About Your EDA Approach (and Do You Want to Disclose Even If It Is Not Required)?*

1. **Meet and Confer Conferences: What Is Required?**

• **What do the rules say?** Both the federal and state rules describe topics for counsel to discuss, but they do not require any particular amount of detail in the discussion. The rules state only that the parties should discuss “any issues” about e-discovery. The conference may address topics such as:
  - A proposed plan and schedule of discovery;

---

8. *See Minn. R. Civ. P. 26.06(b), 26.06(c)(3). The court may direct the attorneys to appear for a discovery conference upon request by a party. Id. R. 26.06(d). The parties must first meet and confer to try to resolve the issues, and the party seeking a conference must file a motion with specific information. Id.*
9. *See id. R. 26.06(d) (2); see also Fed. R. Civ. P. 26(f) (2).*
Any issues relating to disclosure or discovery of electronically stored information, including the form or forms in which it should be produced;\textsuperscript{10}

Any limitations to be placed on discovery.\textsuperscript{11}

- The parties “must consider the nature and basis of their claims and defenses . . . ; discuss any issues about preserving discoverable information; and develop a proposed discovery plan.”\textsuperscript{12}

- The “discovery plan must state the parties’ views” on several things, including:
  - “[T]he subjects on which discovery may be needed, when discovery should be completed, and whether discovery should be conducted in phases or be limited to or focused on particular issues.”\textsuperscript{13}
  - “[A]ny issues about disclosure or discovery of electronically stored information, including the form or forms in which it should be produced.”\textsuperscript{14}
  - “[W]hat changes should be made in the limitations on discovery imposed under these rules.”\textsuperscript{15}

- How are these conferences being conducted? There is wide variation. Some attorneys go into incredible detail and bring IT staff; others see it as a mere formality. However, the increasing trend amongst practitioners is to be well prepared to discuss electronically stored information at such conferences, and some courts have scolded parties for not engaging in a meaningful Rule 26(f) conference.\textsuperscript{16} Some jurisdictions have

\textsuperscript{10} See FED. R. CIV. P. 26(f)(3)(C).


\textsuperscript{12} FED. R. CIV. P. 26(f)(2).

\textsuperscript{13} See id. R. 26(f)(3).

\textsuperscript{14} Id. R. 26(f)(3)(B) (emphasis added).

\textsuperscript{15} Id. R. 26(f)(3)(C) (emphasis added).

\textsuperscript{16} Id. R. 26(f)(3)(E) (emphasis added).

\textsuperscript{17} See, e.g., Hanwha Azdel, Inc. v. C & D Zodiac, Inc., No. 6:12-cv-00023, 2012 WL 6726412, at *1 (W.D. Va. Dec. 27, 2012) (scolding the parties for failing to come up with a “meaningful plan” for ESI discovery and stating that “Rule 26 recognizes the unique problems posed by the discovery of ESI and requires parties cooperate[,] . . . [and] mandates that the parties meet and confer . . . .”); Kleen
created requirements for certain topics that must be addressed during the Rule 26(f) conference. 18 Being knowledgeable about your client’s electronic records and data can prevent significant headaches and the need to “redo” discovery productions. It also enables the parties to stipulate regarding what sources of data will not be produced, or to phase discovery so that the most relevant and fruitful sources of discovery are located, reviewed, and produced first, and time and money spent reviewing and producing data that are expensive to retrieve/review and of minimal value can be limited.

- Aspirational principles: Several sources, including the Sedona Conference, recommend that counsel engage in a thorough Rule 26(f) conference. 19 Both the Minnesota Rules of Civil Procedure and the Federal Rules of Civil Procedure require


counsel to confer on discovery issues.\textsuperscript{20} In the 2006 revisions to Federal Rule 26, the Advisory Committee noted that “[w]hen the parties . . . anticipate disclosure or discovery of electronically stored information, discussion at the outset may avoid difficulties or ease their resolution.”\textsuperscript{21} The Advisory Committee also stated, “Early identification of disputes over the forms of production may help avoid the expense and delay of searches or productions using inappropriate forms.”\textsuperscript{22} As one court noted, “Of course, the best solution in the entire area of electronic discovery is cooperation among counsel.”\textsuperscript{23}

• **Is counsel required to share search terms?** There is no specific requirement that search terms be shared. For the producing party, it may be beneficial to discuss and agree upon search terms. Search terms can narrow the scope of documents to be reviewed, and having the agreement of opposing counsel will help the producing party if it later faces a motion to compel. Counsel should not agree on final search terms until they have done EDA of the search terms to identify false hits or problematic terms.

2. **Dos and Don’ts of Meet-and-Confer Conferences**

**What to Do Before and at Meet-and-Confer Conferences:**

• Come with a list of technical specifications provided by your litigation support team (or bring a team member!).\textsuperscript{24}

• Discuss whether documents will be produced as native files or TIFF/PDFs. This can vary by document type. For example, some people prefer to produce/receive Excel spreadsheets in native format with intact formulae and produce/receive all other documents as TIFF images.\textsuperscript{25}

\textsuperscript{20} See Fed. R. Civ. P. 26(f); Minn. R. Civ. P. 26.06.


\textsuperscript{22} Id.


\textsuperscript{24} See Fed. R. Civ. P. 26, advisory committee’s note (2006 amendment) (“It may be important for the parties to discuss those [electronic storage] systems, and accordingly important for counsel to become familiar with those systems before the conference.”).

\textsuperscript{25} See id. (“Early identification of disputes over the forms of production may
Discuss whether discovery can be conducted in phases, such as starting with a limited number of custodians.

Discuss whether the parties can agree on date ranges for discovery requests.

Perform EDA prior to the conference, if possible. If not possible, perform EDA before engaging opposing counsel in a dispute about the scope of discovery. It is much more effective to be able to support your arguments with hard numbers. For example, “Additional custodians are unnecessary. Limiting the production to four custodians is reasonable because it still yields 60,000 documents.”

Discuss whether you can agree that certain sources of ESI (such as backup tapes) are “not reasonably accessible.”

Discuss a “claw-back” agreement for inadvertently produced documents.

Don’ts of Meet-and-Confer Conferences:

Don’t assume that the discovery conference is a routine event that requires no preparation. This is an opportunity to save time and money for your client, and should be used to your client’s full advantage.

Don’t assume that time spent preparing for the Rule 26(f) conference is wasted. Preparation and planning can be well worth the time and client’s money, particularly time spent identifying technical specifications and coming up with a proposed list of initial custodians. A small amount of time planning can save large amounts of money in the collection, processing, review, and production aspects of the case.

Don’t agree to search terms without running EDA. It may be necessary to agree on tentative search terms, but EDA is necessary to ensure that you do not agree to inappropriate search terms. For example, parties searching for e-mails about a specific type of transaction (e.g., “accounts receivable”) may inadvertently select terms that appear in e-mail signature
blocks (e.g., “Accounts Receivable Manager”), leading to huge numbers of false hits. If the data have already been collected, perform EDA and develop a list of potential terms prior to the Rule 26(f) conference. If collection happens later, follow up with counsel after EDA has been performed to finalize the search terms.

- Don’t ignore the technical requirements. This is usually a simple and noncontroversial topic at the beginning of litigation, but failure to identify technical specifications can lead to huge costs later in the litigation. For example, it can be hugely costly and time consuming to convert data that are in the “wrong” format after the fact.

- Don’t assume that counsel must agree on every issue at the initial Rule 26(f) conference. Counsel may be able to agree on a few search terms and custodians at the initial conference and can confer later if a second phase of discovery is necessary.

- Don’t forget to do anything on the “DO” list!

3. Obligation to Use Adequate Search Terms.

Electronic discovery requires cooperation between opposing counsel and transparency in all aspects of preservation and production of ESI. Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords, with input from the ESI’s custodians as to the words and abbreviations they use, and the proposed

27. Opposing counsel may raise an issue that requires further investigation. For example, opposing counsel may request that a certain custodian be included in the collection and production. It may be necessary to confer with your client to identify that custodian’s role in the facts underlying the litigation before further discussing the issue. If the opposing attorney is mistaken about the employee’s role, he or she may agree to exclude the employee from the collection. Or, the attorney may learn that the employee had a key role, and may agree to collect and produce the documents. Likewise, attorneys may need additional time to further investigate issues such as: unexpected document formats, unexpected storage formats, or appropriateness of search terms. When unanticipated issues arise, or the attorneys are unable to come to an agreement, it may be wise to postpone a final decision and agree to gather additional information and address the issue in a later conversation.
methodology must be quality control tested to assure accuracy in retrieval and elimination of “false positives.” It is time that the Bar—even those lawyers who did not come of age in the computer era—understand this.\(^\text{28}\)

There are very few cases dealing with the adequacy of search terms, but courts are showing increasing interest and sophistication in the document collection and review process.

- **Reasonableness of requested search terms:** Search terms must be reasonable for both parties. They must be both reasonable in number and reasonably calculated to retrieve relevant information (beneficial for the party requesting production), and cannot be unreasonably broad or burdensome (beneficial for the party doing the production).\(^\text{29}\)

- **Expert testimony and competency of court and counsel to determine whether search terms are reasonable:** In extreme cases, it may be necessary to have an expert testify about the reasonableness of search terms.\(^\text{30}\)

- **Use of EDA information to support litigation positions:** A party must be able to defend its selection of search criteria or support its objections to an opponent’s discovery requests.\(^\text{31}\)


\(^{29}\) See id. at 135. Parties must confer and agree on search terms, with input from custodians about likely search terms.

\(^{30}\) See United States v. O’Keefe, 537 F. Supp. 2d 14, 24 (D.D.C. 2008) (“Whether search terms or ‘keywords’ will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics and linguistics . . . . 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. This topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence.” (citation omitted)).

\(^{31}\) See Pension Comm. of Univ. of Montreal Pension Plan v. Banc of Am. Sec., L.L.C., 685 F. Supp. 2d 456, 465 (S.D.N.Y. 2010) (noting a party’s obligation to “assess the accuracy and validity of selected search terms”), abrogated by Chin v. Port Authority, 685 F.3d 135 (2d Cir. 2012); Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251, 262 (D. Md. 2008) (chastising a party for “fail[ing] to demonstrate that the keyword search [it] performed on the text-searchable ESI was reasonable”). A party (and its attorneys) is obligated to either object or respond to properly issued discovery requests under Rule 26. If a party performs an inadequate collection or search, the party may be sanctioned for improper discovery conduct. See, e.g., Qualcomm Inc. v. Broadcom Corp., No. 05cv1958-B
can be used to prove that discovery requests are unduly burdensome when objecting to a motion to compel.\textsuperscript{32}

\textbf{E. Additional Sources of Information Relating to Early Data Assessment}

\textbf{1. Judicial}

- \textsc{Barbara J. Rothstein} et al., \textsc{Managing Discovery of Electronic Information: A Pocket Guide for Judges} (2007).

\textbf{2. Bar Association}


\textbf{3. Academic}

- Michael R. Arkfeld, \textsc{Arkfeld on Electronic Discovery and Evidence} (2d ed. 2007).


\textsc{32. See Victor Stanley, Inc.}, 250 F.R.D. at 260, 262 (“Selection of the appropriate search and information retrieval technique requires careful advance planning . . . . [T]he party selecting the methodology must be prepared to explain the rationale for the method chosen to the court, demonstrate that it is appropriate for the task, and show that it was properly implemented.”).


II. EFFICIENT AND DEFENSIBLE REVIEW

A. Cost of Review

Document review costs continue to soar; they are regularly perceived as the most expensive aspect of conducting litigation in the information age. Too often, this leaves legal stakeholders with the challenge of determining whether or not the cost and burden
of identifying and producing electronically stored information is proportionate to the importance of resolving the issues in dispute.\textsuperscript{33} A manual document review process is the most common means to identify responsive (or privileged) electronically stored information, the cost of which dominates the e-discovery process.\textsuperscript{34}

In a recent RAND report, it was reported that for the cases studied, e-discovery costs ranged from $17,000 to $27 million, with a median of $18 million.\textsuperscript{35} The study went further to estimate that about seventy percent of the cost of e-discovery arises from the review process, which includes a relevancy review often conducted by attorneys and paralegals.\textsuperscript{36}

Much of the cost of document review can be attributed to the volume of documents to be reviewed, the quality of the documents being reviewed, and the actual review process, which will be discussed further in Part II below.

B. Managed Review

Managed review broadly refers to the process of supervising document review performed by a group of attorneys or non-attorney litigation support personnel. Historically, this had been a junior associate attorney at a law firm essentially monitoring either a group of more junior attorneys or a group of attorney contractors, assigning document sets, and answering the review team’s questions as they arose.

In recent years, however, the expansion of ESI has resulted in the potential number of relevant documents exponentially exploding in matters.\textsuperscript{37} Additionally, the type and complexity of

\begin{itemize}
  \item \textsuperscript{34} Id. at *7.
  \item \textsuperscript{35} See Nicholas Pace & Laura Zakaras, Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery 17 (2012).
  \item \textsuperscript{36} See id. at 25.
  \item \textsuperscript{37} Ralph Losey, Rethinking Relevancy: A Call to Change the Rules to Narrow the Scope of ESI Relevancy, E-DISCOVERY TEAM (Jan. 24, 2011, 8:08 PM), http://e-discoveryteam.com/2011/01/24/rethinking-relevancy-a-call-to-change-the-rules
\end{itemize}
technology used to review these documents has expanded exponentially as well. As a result, managed review has evolved into a specialized skill set that is practiced by experienced attorneys in firms, litigation management professionals (who may or may not be attorneys themselves), and specialized companies. These skills involve not only knowledge of the law and technology, but also logic and process design. When used effectively, these skills allow for the creation of workflows that are quick, cost effective, and accurate at reviewing and coding documents.

1. Review Workflow

Before commencing a managed review project, the workflow should be planned to lay out the path that documents will follow from start to finish. The larger the project size, the more likely that many phases of the workflow will need to be ongoing at the same time so that the project can be completed on a timely basis.

As early as practicable before commencing the managed review, the client (including client IT representatives), outside counsel, database vendor, and legal services company (if applicable) should have a meeting to discuss roles and responsibilities. These parties should all agree on a workflow to assure the managed review is completed in time and that each document is properly sent through each step in the workflow.

The workflow in a managed review can often be broken into the following steps:

- **Early case assessment.** Depending on the circumstances of the review, ECA can be a valuable tool to gain high-level information about the documents in a case, gather more documents from the client, and cull documents already gathered to exclude nonrelevant documents. ECA is often facilitated by the database provider with a specialized tool. Though ECA tools vary between database providers, they may allow users to view and sort the documents based on e-mail participants, file types, dates, concept clusters, and keywords.

- ECA may also be used in conjunction with a limited review. For example, a sample set of data may be loaded for review for testing proposed keyword search terms to estimate the
number of potentially responsive documents and the accuracy of the terms in identifying responsive documents. Outside counsel or a small group of skilled contract attorneys may review sample sets of documents. Based on the results, changes may be made to search terms and tested to improve the accuracy of keywords before documents are processed for first review to help reduce overall processing and review costs.

- **First review.** Once the documents that will be reviewed as part of the project are determined, the process of review can begin. The First Review team will generally be made up of attorneys dedicated to this process, such as contract attorneys who are specifically retained for the review project and whose involvement in the case will terminate once the managed review is completed. These attorneys should be asked to complete a conflict of interest form listing any potential connections with any litigant or counsel in the matter. A contractual nondisclosure agreement may also be executed for added confidentiality. Additionally, many managed review providers also perform periodic background checks and bar status verifications.

- The size of the First Review team will typically be determined by the number of documents that need to be completed and the deadline for the project. The number of attorneys can vary from two to 100 or more. In order to keep consistency and leverage knowledge developed by the team over the course of the project, the team should be kept as small as possible to meet the necessary deadline.

- **Quality control:** Every managed review should have a team dedicated to quality control to catch mistakes made by the First Review team as well as to apply knowledge developed later in the review to previously reviewed documents. As with the First Review team, the Quality Control team will be made of attorney contractors who often have been retained solely for the managed review portion of the project. Typically, this team

---

38. Because a lawyer owes a duty of undivided loyalty to a client, ethics rules apply to conflicts of interest. See Model Rules of Prof’l Conduct R. 1.7–1.10, 1.13 (2011). In light of this ethical imperative, it is a recommended practice that a screening occurs to avoid impermissible conflicts of interest.
should be made up of the more senior attorney contractors and about fifteen to twenty percent of the size of the First Review team.

- **Quality control team’s review**: The Quality Control team should perform first review for an initial phase of the project to become familiar with the documents in the review. Quality control should begin and be performed while the First Review team is still reviewing to provide feedback to the First Review team as close to real time as possible to prevent mistakes before additional documents are coded in the same manner.

- **Privilege review**: If the managed review project requires the creation of a privilege log, then a separate privilege team is useful to draft such logs. This team should be kept as small as practicable to improve consistency and may be a subset of the Quality Control team. Moreover, because of the nature of the subject matter and importance of this step, in particular, this review ordinarily should not be delegated to nonlawyers.

- **Counsel review (second level review)**: A final review of the documents should be performed by the attorneys permanently assigned to the matter, whether in-house or outside counsel. This should include review of documents identified as privileged to make final determinations of what should be withheld. This review should also include sampling of responsive and nonresponsive documents to assure that the project teams are applying coding properly.

  - Managed review can utilize attorneys of all skill levels. The First Review team can comprise attorneys at any experience level. More senior attorney contractors with significant prior managed review experience should be used on the Quality Control or Privilege Review teams. For specialized matters, such as patent litigation, efforts can be made to recruit attorneys with experience in particular areas of law.

2. **Prioritization of Review**

The advent of sophisticated databases to manage stored ESI allows for the prioritization of documents that meet specific criteria in the managed review process. Documents can be organized using any or all of the following categories to help speed up the identification of the most important documents and the review process overall.
• **By custodian:** Somewhat of a holdover from paper document reviews, organizing documents by custodian can still be useful for matters where the managed review process may be ongoing and allow for the production of the complete set of documents for particular custodians. This can be helpful if the managed review is running contemporaneous with an oral discovery or a rolling production schedule.

• **Potentially privileged documents:** The managed review of privileged documents can take longer to complete. This is because these documents are likely be reviewed by a quality control process as well as a final review by the counsel team for the matter. Additionally, documents deemed privileged will typically need to be cataloged for disclosure on a privilege log. Using searches to pull documents that are potentially privileged to the beginning of the managed review process will assure these documents have the most time available to complete these additional steps.

• **Responsive terms or concepts:** One of the primary goals of the managed review process is the quick identification of potentially important documents related to the matter at issue. The sooner these documents can be identified, the more time in-house and outside counsel will have to use them in their assessment of the case. To that end, parties engaged in the managed review process should make an effort to search out and review documents containing multiple responsive terms or concepts earlier in the review process.

3. **Predictive Coding**

Predictive coding is a form of automated document review that involves a combination of people, process, and technology, whereby the technology “learns” which documents may be relevant by analyzing a small group of documents that attorneys have manually reviewed and coded. The process and methodology can vary by case and by the type of technology used for the automated review.

Predictive coding does not mean there is no human review of responsive documents. Predictive coding often begins with a human expert or small team of experts familiar with the collected documents. The subject matter expert(s) review randomly selected documents from what has been collected, or review chosen “seed sets”\(^\text{40}\) of documents to determine which documents are responsive or relevant, and which are nonresponsive (and perhaps also for privilege). Different technologies are then used to categorize or rank the remaining documents as relevant or not relevant. Quality assurance checks are then conducted by attorneys or other people familiar with the case to determine the accuracy of finding relevant documents, which allows for a subsequent “smarter” automated search if necessary. The quality assurance checks often include a human review of a sample of the documents categorized as relevant as well as a sample of those categorized as irrelevant.

The hope is that automated review will reduce discovery costs. Proponents of predictive coding contend that it can be more accurate than traditional human review of all collected documents or other traditional forms of technology-assisted review (such as keyword searches) and result in production of more relevant documents at an earlier stage in the litigation.\(^\text{41}\)

\(\text{a. Considerations}\)

Transparency is crucial, and an agreement with opposing counsel or court approval in the early stages of a case as to methodology is as well. Agreement is more likely when opposing counsel and a judge understand the methodology of the automated review, including quality control. Topics of negotiation with opposing counsel and topics a court will likely want to discuss include statistical sampling methods, confidence levels, and precision/recall.

---

\(^{40}\) “Seed set” is the initial sample of documents coded by one or more subject matter experts as relevant or not relevant provided to teach a Machine Learning Algorithm how to distinguish between relevant and not relevant documents beyond those in the seed set. See Maura R. Grossman & Gordon V. Cormack, The Grossman-Cormack Glossary of Technology Assisted Review, 7 FED. CTNS. L. REV. 1, 29 (2013).

\(^{41}\) Kerschberg, supra note 39.
Seed set identification used to train or educate the computer program/software is important. Consider ahead of time whether you are willing to share the seed sets with opposing counsel, including those documents identified as nonresponsive. Take care to not produce potentially privileged documents in the seed sets if the seed sets will be shared with opposing counsel. A claw-back agreement with opposing counsel or a court order requiring inadvertently produced privileged documents be returned is recommended and should be in place ahead of any production to minimize the risk of otherwise privileged documents being used against the client.

If the documents to be reviewed include ones in a foreign language, determine if the tool selected is useful for review of those foreign language documents.

b. Potential Drawbacks of Predictive Coding

Predictive coding tools might be ineffective with some document types, such as image-based files, audio files, and perhaps Excel spreadsheets or other documents containing mostly numbers. The nature of the content of those documents makes it difficult for a machine learning algorithm to learn whether such documents are or are not responsive or relevant.

Predictive coding may also not ultimately result in cost savings. It might be more expensive up front with attorney time in preliminary review/training. Furthermore, in a phased discovery approach, or with unique custodian groups, there can be significant retraining requirements.

There is fear about relying on a computer algorithm to determine what is responsive, nonresponsive, and privileged; risks are always inherent in using evolving technology. Using larger seed sets and increasing the number of human-performed quality assurance checks of computer algorithm decisions might increase

confidence in the reliability of predictive coding, but also would tend to diminish the extent of any advantage gained by employing predictive coding.

4. Documentation

Repeatability and traceability through documentation of workflows and all guidelines associated with the review is essential in the event that the results of a review are later challenged. There are several reasons the result of a managed review could be challenged, including a motion to compel claiming a party has failed to identify all responsive documents, or as part of an inadvertent production claw-back effort to demonstrate the efforts used to prevent disclosure of privileged material. Because these challenges may occur months after the review of documents has concluded, thorough documentation can help recreate the methods used to perform the managed review long after it has been completed.

Though a more thorough description of which documents are necessary for an efficient managed review is described below, documentation of the managed review process should include the following materials:

- **Review training materials:** A copy of the training materials used as part of the project should be kept—such as a PowerPoint or other presentation and speaker notes. A log should also be kept of the dates the materials were presented, who presented the material, and who was present at the training.

- **Review binder materials:** A complete copy of the review binders that were provided to each reviewer in either paper or electronic form should be kept. Also, a log sheet showing the checkout and check-in of printed binder materials to attorneys should be kept as well. Additionally, if content is updated

---

43. *See supra* notes 31–32 and accompanying text.

44. “When a producing party claims inadvertent disclosure, it has the burden of proving that the disclosure was truly inadvertent.” Fox v. Massey-Ferguson, Inc., 172 F.R.D. 653, 671 (E.D. Mich. 1995). *See, e.g.,* Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251, 260–68 (D. Md. 2008) (holding that the attorney-client privilege did not protect a number of documents the defendant claimed it inadvertently produced to the plaintiff).

45. *See infra* Part II.B.3.
during the review, copies of the original materials should be kept and the date of the updates logged.

- **Log of review issues and resolutions**: A complete log showing the questions raised by the review team and the answers provided by the attorneys supervising the project should also be kept. This log should contain the date of the question and an identification number for any example documents.

If questions are asked and answered during a meeting with counsel and the review team, they should also be added to the written log. Additionally, the log should contain any changes in coding instructions, additions of issue codes, and privileged persons discovered in the documents and the date thereof, among other things.

This log is valuable to demonstrate counsel’s oversight of the review process, and can also help to demonstrate why documents may have been coded one way earlier in the managed review and a different way later on.

5. **Training the Review Team**

The counsel on the matter should provide substantive training to the members of the review team. Additional training should be provided to the team on the review platform being used for the managed review project.

This training may take place in person, via videoconference, Internet meeting, or even conference call, depending on the locations of counsel, the review team, and available facilities/technology. The training should contain background information on the purpose of the review to help provide context to the review team. This background information can help the review team spot important issues that even counsel may be unaware of at the outset of the project. The training should also contain detailed instructions on how documents should be coded. These instructions should be as objective as possible to ensure consistent application of the coding instructions across the review team.

When possible, the review team should be provided with the following materials as part of their training on the matter:

- **Coding protocol**: This document should contain background information on the matter and should lay out the criteria to be used in coding the documents in as objective terms as possible.
These criteria should describe what kinds of documents are responsive to the litigation, as well as define any issue codes the review team should apply to the documents they review. This document should also contain a date or version number so that updates and changes can be distinguished. If additions or changes to the coding criteria are made during the review, they should be reflected in an updated version of this document when possible.

- **List of privileged persons**: This document should contain a list of any persons that the review team should consider as a person whose communications should potentially be protected by the attorney-client privilege or work product doctrine. If possible, this list should be in an electronic format to enable easy searching by the review team and to allow for easy updating for additional persons that may be discovered during the review.

- **List of persons likely to appear in documents**: This document can be created by outside or in-house counsel and should list persons who are likely to appear as authors or recipients of the documents subject to review. Any relevant background information that is known about these people, such as their known or likely roles in the events at issue and their positions at the various companies during the relevant time period, should be included to provide context to the review team. An organizational chart is very helpful as well.

- **List of acronyms or technical words likely to appear in documents**: This document can be created by outside or in-house counsel and is very helpful for highly technical or jargon-laden documents. A list providing definitions for frequently used acronyms or technical words can help to speed review rates by allowing the quick determination of meanings. If necessary, the review team can create such a list as the review project progresses. A list created during the review can provide counsel with valuable information as well. This list should also be kept in an electronic format so searches can be easily updated.

- **Examples of priority documents**: Before a review begins, typically several key documents have already been identified, by either the document custodians themselves or by counsel, as highly important documents. In reviews where the team is being asked to mark documents as “priority,” providing examples of
these types of documents to the review team can help prevent them from over- or under-marking documents as priority.

- **Examples of “close call” documents:** Often there may be documents that appear nonresponsive that should actually be considered responsive, and vice versa. To the extent that examples of such documents can be identified prior to the review training, they should be shared with the review team to help prevent documents from being incorrectly coded early on in the review.

Once initial training is completed, counsel should plan to either be on site to answer initial questions from the review team or to have a follow-up meeting to address questions the following day. Meetings should continue to be held to resolve questions and advise counsel of issues in the documents; the frequency of these meeting may reduce as the team becomes more comfortable with the material.

### 6. Quality Control

Quality control is the review of an already-coded document to verify that it was marked correctly on first review. Excellent quality control is imperative for a successful managed review process because mistakes on documents during the first review phase are certain to occur, and because changes to coding instructions often occur during the course of any document review. The goal of the quality control process should be the identification and correction of mistakenly coded documents and documents that were coded before counsel communicated any change to the coding instructions.

To this end, the quality control process should be focused on reducing errors in the coded documents rather than performing a second check on a certain percentage of documents. As a general reference, a properly conducted quality control review usually entails re-reviewing fifteen to twenty-five percent of all documents coded by the team. Documents should be selected based on their likelihood of a coding mistake. The following categories of documents should be included in quality control checks:

- **Documents with conflicting coding:** Though some review platforms will allow the coding template to be set up in such a way as to prevent many coding mistakes from occurring, any documents that are marked in such a way that the coding is incomplete or
plainly incorrect (such as documents marked nonresponsive with an issue code or marked responsive with no issue code identified) must be re-reviewed by the Quality Control team to remove those errors.

- **Potentially privileged documents not marked privileged**: Any responsive documents that contain potentially privileged terms, but not marked privileged on first review, must be re-reviewed by the Quality Control team to minimize the chances that a privileged document is produced.

- **Nonresponsive documents containing responsive terms**: Documents that contain a high number of terms usually found in responsive documents, but were marked nonresponsive on first review, must be re-reviewed by the Quality Control team to correct any omissions from production.

Quality control review of a managed review project should be launched after the Quality Control team has had the opportunity to code documents on first review and to learn about the issues in the documents. Quality control review should take place contemporaneously with first review so that the quality control team can provide real time feedback to counsel the first review team about mistakes being found in documents. Because documents cannot be selected for quality control until they have been through first review, depending on the size of the review, the quality control team may not finish its work until days or weeks after the first review team has completed its review. Accordingly, time for the completion of the quality control process should be built into the total timeline for the managed review project.

7. **Productivity Tracking and Review Metrics**

The use of sophisticated, dedicated document review platforms instead of paper reviews or reviews conducted on nondedicated databases, such as Summation, has resulted in the development of specific review productivity and quality tracking metrics. These metrics can be available directly within a review platform or through some third-party add-on service.

These metric platforms will allow for the tracking of individual user productivity and team totals. Monitoring individual user productivity allows for the identification of team members that may be reviewing too quickly or too slowly. These tracking metrics often contain “what if” capabilities that allow forecasting based on
different variables, such as deadline date, team size, and team productivity rate. This allows the manager to see if the project is on track to finish by a required date or if adjustments to the team members or review rates will be necessary.

Additionally, these metrics can improve quality by tracking error rates in documents based on quality control. This error rate tracking can help spot-review attorneys that may need additional training on an aspect of the review or issue codes that the entire team may need clarification or additional training on applying.

8. Pricing Models

Currently there are two primary ways that managed review services are priced: hourly pricing and project-based pricing. Hourly pricing has historically been used for these services, but similar to other legal services, clients face an open-ended expense and there is little incentive on the service provider’s part to work to minimize the number of total hours billed. Accordingly, clients have increasingly requested these services on a project-based model of a price per unit to help make costs predictable. Due to their varying complexity and time intensity, some services, such as redactions and privilege log drafting, do not lend themselves to a predictable cost model and may still be done on an hourly basis.

- **Hourly-based pricing model**: Historically, managed review services have been provided on an hourly basis with fees based on the skill level of the attorney contractors used. Typically, these fees are lower for first-level review team members, higher for senior attorney contractors or those with special skills (such as

46. Metrics measured might include hours worked by the reviewers, hours logged on the review platform by the reviewers, average hours worked and logged, average number of documents/pages coded, number or percentage of documents checked for accuracy (quality control), error rates, system downtime, and number or percentage of documents not yet coded. Using Metrics in E-Discovery, EDRM, www.edrm.net/resources/standards/edrm-metrics/edrm-metrics-case-study (last visited Oct. 28, 2013).

47. Issue codes are categories established for document reviewers to classify documents based upon subject matter to allow documents to be organized, analyzed, and retrieved more easily later in the litigation. Sally Kane, Document Coder, ABOUT.COM, http://legalcareers.about.com/od/entrylevellegalcareers/p/Documentcoder.htm (last visited Dec. 18, 2013).
foreign languages), and include a management-level fee to oversee the project.

- **Project-based pricing model**: Similar to other legal services, clients are requesting that services be provided on a project basis to allow the client to accurately forecast the cost of the review. Formats for this pricing are usually a price per document or price per page. This incentivizes the managed review provider to make sure that documents are reviewed efficiently, in terms of time and resources, and accurately to minimize the amount of re-review to correct erroneous coding.

C. **Roles and Responsibilities**

For document review projects, competent project management is important. All projects will benefit from clear and consistent management structure. “Applied to the process of e-discovery, it is essential that the role of project leader be clearly and decisively vested in one or more individuals who are empowered by the client to manage the effort of counsel and service providers.”

1. **Client**

Client responsibilities may vary from matter to matter depending on time and resources dedicated to full-time management of the e-discovery process. While in-house counsel plays an important role in managing the litigation and investigation, active supervision of e-discovery on a daily basis may not be practical. In-house counsel should work with their law firm/outside counsel and the review vendor to ensure appropriate case and project managers are assigned to oversee the entire process.

Ultimately, all legal counsel, both in-house and outside counsel, have a duty to supervise document productions by ensuring that reasonable steps are taken so that document productions are accurate and complete. The client should allow


the management teams charged with overseeing the document review sufficient time and resources to supervise the collection, review, and production process.

2. Law Firm

As with in-house counsel, outside counsel representing the matter must take necessary steps to ensure that document productions are accurate and complete. This ultimately means outside counsel must supervise the document review and production process. The most effective means to facilitate supervision over the entire process is for outside counsel to assign a case manager that will work with the client and review vendor to oversee the matter.

The role of case manager is an important one, typically filled by outside counsel. As the case manager, the individual will work with the client and vendors to define the project’s budgets, goals, and objectives and develop a plan of execution. The case manager will understand both the substantive and strategic aspects of the litigation, while also having experience in the various phases of e-discovery.

Ultimately, the case manager, or other outside counsel managing the litigation, will be required by court or agency rules to certify the discovery responses, which can result in consequences if a challenge is made against the accuracy or diligence of the discovery efforts. Thus, it is important that the case manager oversees and communicates throughout the entire discovery process, including those parts handled by the client or the vendor, or both.

Lawyers and client organizations can be sanctioned for any perceived shortcomings or failures in the discovery process, making effective project management key.

50. See id.
52. See id. (citing Fed. R. Civ. P. 26(g)).
3. Review Vendor

At the outset of a matter, the review vendor will designate a project manager to supervise the day-to-day activities of the review team, manage the quality control process, and triage any issues that arise. The review vendor may also utilize a lead reviewer, or multiple lead reviewers depending on the matter’s scope, that can be available to proactively answer questions for the review team and address any issues as they arise. The lead reviewer serves as a liaison between the review team and the main project manager overseeing the entire review project.

Effective management by the project manager will include weekly or biweekly status calls with the law firm case manager or the client, ensuring that an appropriate quality control occurs throughout the life of the review, managing the pace of the review project to ensure meeting production deadlines and that the review team is adequately staffed and receiving review materials, preparing the privilege log, and overseeing the redaction process.

4. Additional Sources of Information on Efficient and Defensible Review


documents, failure to produce other documents, and failure to preserve relevant evidence).
III. PRODUCTION FORMAT

A. Production Specifications

Begin planning for production early in a case, including determining what database or document review tools you will be using. The lawyers should provide detailed, written instructions to opposing counsel to outline your specifications. As a general rule, do not request a specification from an opponent that you do not want to provide. Remember that the rules require that the form of production be “reasonably usable.”

B. Standard Production Requests

These formats will change as technology changes. As of 2013, the following are standard production request formats:

- Load files;
- Image files—TIFF, PDF, JPEGs, or native files;
- Production of Excel and PowerPoint documents in native format;
- Optical character recognition (OCR) for hard copy;
- Extracted text for electronic documents;
- Metadata fields.

C. Production of Specific File Types

1. Native Format

Some file types are best produced in native format because they do not convert well. For instance, in 2013, Excel files do not convert well and are typically produced in native format. The same is true for the speaker notes in PowerPoint presentations.

54. For more information on production formats, see Appendix A, infra.
2. **OCR/ICR**

Other files are produced with OCR/ICR, which are optical character recognition/intelligent character recognition. This allows the software program to look at the pages of the documents and read the characters, allowing for full text searching of content. Remember that with OCR and ICR, the reading is only as good as the original document (contingent upon how well the software can read the document). These programs are not 100% accurate. They can be used on files that are not text based and also on hard-copy scans, digital photographs, and other image files.

3. **Extracted Text**

Extracted text applies to documents only. When the documents are processed, the text is extracted and put into a load file. Text extraction allows for full text searching and is 100% accurate, thus eliminating the need for OCR.

4. **Metadata**

Metadata answers the “who, what, when, where, and how” about every piece of data. It is the data behind a document—who created it, the date of creation, when it was last modified, tracked changes, etc. There can be over 1000 metadata fields for every document. Parties should discuss and agree on what fields are going to be produced. Metadata can be a valuable litigation and internal investigation tool because it details a document’s history and distribution. Metadata can help reconstruct a timeline of events, produce additional leads for investigation, and establish a person’s knowledge of the existence and content of files. Examples of the standard metadata fields are as follows: author, recipient, date created, date(s) modified, title or e-mail subject, file name, carbon copies and blind carbon copies, attachments, source or

---

custodian, date sent, date received, last date accessed, and file size. Do you have to produce metadata? The simple answer is YES!

If the data are maintained electronically, they cannot be produced in a form that removes or degrades the receiving party’s ability to use or search the data. Production of “an electronic document in the form in which it is regularly maintained . . . must include all metadata . . . ”

D. Final Thoughts About Production

Form of production should be discussed early in a case. It is important to communicate with opposing counsel to agree on how the electronic information will be produced to each side. This saves time and money by ensuring information is produced in a format allowing for an efficient review of what is produced and minimizes any possibility of having to produce any of the same information a second time in a different format. Negotiating ahead of time minimizes the risk of producing information in a format a court might later deem to be unacceptable. If agreement cannot be reached with opposing counsel, seeking early judicial intervention also results in a timelier (and hopefully helpful) judicial guidance.

All parties should provide detailed, written production specifications to the other side to better ensure that the documents received are in a format compatible with the document review tool being used to review the opponent’s production. It is important to understand what your opponent is requesting for production format specifications and to not agree to a request until you understand it.

Planning ahead is helpful. Knowing how you plan to review documents and what tool you will use for review allows you to make certain that the form of production is compatible with the planned method of review.

IV. Social Media and E-Discovery

This section addresses certain aspects of social media that might generate information called for in response to discovery. Social media is a form of electronic communication “through which users create online communities to share information, ideas, personal messages, [videos,] and other content.”

A. Social Media and ESI

Social media is not a passing trend. It is used by billions of people to communicate and share information. In a single day, YouTube users upload twelve years of video, Instagram users upload forty million photos, Facebook users share 2.5 billion pieces of content, and Twitter users send 400 million tweets.

1. Types of Social Media Content

- Friends, Friends of Friends, Connections, Followers, etc.;
- Status Updates, Relationship Status;
- E-mails, Chats, Text Messages, Friend Requests, Pokes;
- Timeline (Profile)—Name, Picture, Gender, Contact, Birthday;
- Wall, Posts, Comments, Tags;
- Likes, Reads, Views, Listens, etc.;
- Networks, Groups, Events;
- Photos, Videos, Audio, Music;
- Apps, App Data, Games.

2. Pushed Content

- E-mail Notifications with Metadata;
- RSS Feeds with Metadata.

60. For information on how electronic discovery differs from digital forensics, see Appendix B, infra.
3. Metadata

- Site Names, Uniform Resource Locators (URLs);
- Date/Time Stamps, Geolocation Information (Check-ins);
- IP Logs, Login/Logout Logs.

B. Potential Social Media Evidence Uses (i.e., Why You Might Want Access to Your Adversary’s Social Media Content).

People share content on social media sites for a variety of reasons. Motivations to share include a desire “to bring valuable or entertaining content to others[,] to define ourselves to others[,] to grow and nourish relationships,” or to publicize things one cares about. What a person or organization shares on a social media site might include information relevant to claims or defenses in litigation matters. Potential evidentiary uses include:

- **Admissions or state of mind**: A social media user who disseminates information to others may not realize that what is disseminated may be used in court against the user who disseminated the information. This includes statements deemed a confession or admission.

- **Witness credibility**: Posts, e-mails, places, friends, and contact information contained on social media sites could impeach testimony that differs from what is posted.

- **Witness character**: Photos, videos, apps, or other information posted by a user might yield critical character evidence which


64. See United States v. Meregildo, 883 F. Supp. 2d 523, 525–26 (S.D.N.Y. 2012) (“Where Facebook privacy settings allow viewership of postings by ‘friends,’ the Government may access them through a cooperating witness who is a ‘friend’ without violating the Fourth Amendment.”).

65. See, e.g., Blayde v. Harrah’s Entertainment, Inc., No. 2:08-cv-02798-BBD-cgc, 2010 WL 5387486, at *5 (W.D. Tenn. Dec. 17, 2010) (using the LinkedIn page of a defense witness identifying his employer as Harrah’s to rebut his trial testimony that Harrah’s was not his employer); Clark v. State, 915 N.E.2d 126, 130 (Ind. 2009) (affirming defendant’s murder conviction and admission of defendant’s My Space posting used to impeach defendant’s trial testimony as to his intent and state of mind).
might be admissible for certain purposes such as criminal sentencing.\footnote{See Eric Tucker, Facebook Used as Character Evidence, Lands Some in Jail, USA TODAY (July 16, 2008, 9:02 PM), http://usatoday30.usatoday.com/tech/webguide/internetlife/2008-07-19-facebook-trials_N.htm.}

C. Discoverability of Social Media Posts

Are social media posts discoverable? Often, the answer is yes, regardless of privacy settings or controls, because the content of social networking sites is not protected from discovery merely because a party deems the content “private.”\footnote{See EEOC v. Simply Storage Mgmt., L.L.C., 270 F.R.D. 430, 434 (S.D. Ind. 2010); see also Glazer v. Fireman’s Fund. Ins. Co., No. 11 Civ. 4374(PGG)(FM), 2012 WL 1197167, at *3 (S.D.N.Y. Apr. 5, 2012); Mackelprang v. Fid. Nat’l Title Agency of Nev., Inc., No. 2:06-cv-00788-JCM-GWF, 2007 WL 119149, at *8 (D. Nev. Jan. 9, 2007) (allowing for the “limited request of production of relevant” private communication from social networks).}


Courts requiring such a showing do so, at least in part, to guard against the “proverbial fishing expedition.”\footnote{Tompkins, 278 F.R.D. at 388; see also Kregg v. Maldonado, 951 N.Y.S.2d 301, 302 (App. Div. 2012) (denying “motion seeking disclosure of all social media account records” as overbroad).} Absent some “threshold showing that the requested information is...
reasonably calculated to lead to the discovery of admissible evidence,” a “[d]efendant would be allowed to engage in the proverbial fishing expedition, in the hope that there might be something of relevance in [p]laintiff’s Facebook account.”

Whether privacy concerns will be “outweighed” depends on relevance, probative value, and danger of unfair prejudice. Courts have restricted what social media content can be produced based on these factors.

D. Requests for Production

1. Social Media Is Discoverable

Social media evidence is ESI and can be discoverable if relevant and accessible. In addition, any party can place a litigation hold on social media by sending a preservation letter to the appropriate Internet service provider for evidence protection. It is also possible to obtain a court order compelling the production of the opposition’s social media evidence. Another strategy is to move to compel a signed consent release from the subscriber or opposing party.


72. See, e.g., Trail v. Lesko, No. GD-10-017249, 2012 WL 2864004 (Pa. Ct. Com. Pl. July 5, 2012) (denying both the plaintiff’s and the defendant’s motion to compel access to the other’s Facebook pages as unreasonably intrusive because, under PA. R. CIV. P. 4011(b), and in this particular case, “the intrusions that such discovery would cause were not offset by any showing that the discovery would assist the requesting party in presenting its case”).

73. See EEOC v. Simply Storage Mgmt., L.L.C., 270 F.R.D. 430, 434–35 (S.D. Ind. 2010) (holding that production of a portion of an employees’ social networking site was appropriate).

74. See O’Grady v. Superior Court, 44 Cal. Rptr. 3d 72, 87 (Ct. App. 2006) (finding that one purpose of the Stored Communications Act is to shield private electronic forms of communication from government intrusion); Flagg v. City of Detroit, 252 F.R.D. 346, 363 (E.D. Mich. 2008) (holding that the Stored Communications Act did not preclude civil discovery of the city’s relevant, nonprivileged electronically stored communications that were maintained by a nonparty service provider, but remained within the city’s control).
2. Strategies in Cases Where the Opposing Party’s Production Is Insufficient

If the opposing party’s production is insufficient because social media evidence is missing or you suspect that evidence may have been destroyed, you can look outside the social media site for information, including e-mail notifications and Really Simple Syndication (RSS) containing content and time stamps pushed out by social media.\textsuperscript{75}

You can also consider whether to move for a court order for computer forensic analysis of the witnesses’ hard drives to recover social media evidence. Courts, however, may be reluctant to order the entire hard drive be produced to an opponent.\textsuperscript{76} For courts sensitive to privacy concerns or protection of confidential business information, claims of “suspected” spoliation must be more than rank speculation and innuendo.\textsuperscript{77} For this reason, the use of a third-party neutral is preferred to handing over an entire hard drive to the opponent’s partisan expert, because the neutral can conduct the forensic analysis, prevent the production of irrelevant or privileged material, and put a stop to fishing expeditions.\textsuperscript{78}

\textsuperscript{75} See RSS (Real Simple Syndication)—Frequently Asked Questions, NEWSCIENTIST, www.newscientist.com/info/in180 (last visited Oct. 28, 2013). RSS content is also referred to as feed and a feed is simply a way in which a reader may subscribe to website content, such as a blog or news site. A news site, for example, may list the latest headlines or entire articles in their feed every time a new article is published. A blog could publish a feed that contains a series of recent posts . . . . Using RSS can consolidate many data sources and stop the need for you to constantly visit many different sites . . . .


\textsuperscript{77} See United States v. O’Keefe, 537 F. Supp. 2d 14, 22 (D.D.C. 2008) (“[V]ague notions that there should have been more than what was produced are speculative and are an insufficient premise for judicial action.”).

\textsuperscript{78} In re Weekley Homes, L.P., 295 S.W.3d at 318.
E. Recovery of Social Media Evidence

Basic information that can be recovered includes:

- User profile (timeline) information (e.g., user contact information, interests, and groups);
- Wall (timeline) posts and content that the user and the user’s friends have posted to his or her profile (timeline);
- Photos and videos that the user has uploaded to his or her account;
- The user’s friend list;
- Notes the user has created;
- Events to which the user has RSVP’d;
- The user’s sent and received messages; and
- Any comments that the user and the user’s friends have made on Wall (timeline) posts, photos, and other profile (timeline) content.\(^79\)

A social media user may collect information, document his or her movements, and disseminate data. Retrieving this information allows one to analyze the user’s pages visited, time and frequency of access, and communications. More advanced information includes the following:

- **IP addresses**: Any IP addresses the user has stored (this won’t be all of the IP addresses that have ever accessed a user account).
- **Login info**: A list of the logins the user has stored (this won’t include every login during the user’s account’s history).
- **Log out info**: The IP addresses from which the user logged out
- **Pending friend requests**: Friend requests the user sent and friend requests the user received but hasn’t accepted or denied.
- **Account status changes**: Dates when the user account was reactivated, deactivated, disabled, or deleted.
- **Poke info**: Information about the pokes the user has exchanged.
- **Events info**: Events the user accepted, declined, and responded “maybe” to.\(^80\)
- **Other profile (timeline) info**.\(^81\)

---


\(^80\). See id.

\(^81\).
The mobile phone numbers user has added to account;
User’s city and hometown (whatever is currently listed);
The names of the family members user has listed on account;
User’s relationship info (names and statuses);
A list of the languages user has added to profile;
A history of any changes user has made to the name on account.

F. Computer Hard Drive Data Recovery

Information that can be recovered from a hard drive includes:
- Facebook status updates, wall posts, and comments;
- LinkedIn search history;
- E-mails and e-mail fragments;
- Chat messages;
- Webpage fragments;
- MySpace live chat;
- Search history;
- Twitter status artifacts;
- Browser history, bookmarks, cookies, icons, logins, autofills, profiles, and other artifacts;
- Peer-to-peer file sharing data;
- Documents, spreadsheets, and images;
- History of connected devices, including portable media;
- History of files opened, sometimes including videos viewed;
- And much, much more.

G. Social Media Visualizers

Sometimes when you are explaining the relationships between people, it is useful to use a visual aid. With technology, it is possible to generate an animated, clickable, visual map of connected

81. Id.
82. If discovery is sought from a party, a request for production of this information directly from that party is the best method to produce it. That party can either hand it over or request his/her own records from the service provider. If discovery is sought from someone who is not a party, a subpoena seeking this information will need to be served on that nonparty. See infra Appendix C.
“friends” from social media profiles. You can then search the visual map to learn how and why people interact, including:

• Which friends know each other;
• Filter social network based on gender or relationship status;
• Discover commonalities, such as group membership and common interests;
• Reveal connections between friends with most photos taken together.

The issue here is relevancy. It is highly improbable that an individual’s entire “map of friends” is relevant to the litigation. But this analysis may still be helpful for an attorney working with a client, but the attorney should ask what utility, if any, this could have as demonstrative evidence.

V. SMART PHONES AND TABLETS: ESI FOR REVIEW

A. Types of Content on Portable Devices

It is now commonplace to be in a meeting where at least one person is taking notes on an electronic tablet. Likewise, people now send or receive both work and personal e-mails on personally owned or company-owned smart phones. Employees and consumers are demanding more flexibility about where they work, how they work, as well as the various devices they use to create, view, and disseminate content. These demands, combined with the increasing technical capabilities of small mobile devices, result in portable devices that may contain a treasure trove of information. Portable devices (phones, tablets, and other devices) can contain a lot of information. The most common types of information are:

• Phone address book (Contacts);
• Appointments and calendar;
• Dialed, received, and missed call logs;
• Text messages (SMS);
• E-mail and attachments;

83. FED. R. EVID. 403; MINN. R. EVID. 403.

84. For information on how electronic discovery differs from digital forensics, see Appendix B, infra.
- Electronic documents;
- Photographs;
- Audio and video recordings;
- Voice memos;
- Multimedia messages (MMS);
- Instant messaging and chat;
- Web browsing history, bookmarks, cookies, etc.;
- Social media (Facebook, Twitter, LinkedIn, MySpace);
- Apps and app data.

B. Other Content—Metadata

In addition to the basic information that a user may store on the portable device, the device itself contains certain information:

- Phone or tablet
  - Make, model, equipment IDs, phone number, etc.;
  - Software versions, language;
  - Date, time, time zone, daylight savings time.

- Forensic tool
  - Identification (make, model, serial number);
  - Software versions;
  - Exam date, time, time zone, daylight savings time.

- Phone or tablet content
  - Hash codes (MD5, SHA1);
  - Date and time stamps;
  - Geolocation information (Geotags);
  - Exchangeable information file format (Exif) data from onboard camera snapshots and video;
  - Access point data from Wi-Fi logins and activity;
  - Reminders.

C. Another Data Source—Service Provider Business Records

In order to retrieve information from the service provider, a party often seeks to serve a subpoena on the custodian of records for the service provider. One of the best sources for finding the custodian of records’ contact information for social media sites, Internet service providers, and phone companies is the website

Pramas et al.: Using Technology to Facilitate Production of E-discovery

Published by Mitchell Hamline Open Access, 2014
It maintains a listing of hundreds of addresses and phone numbers for custodians of records.

The Electronic Communications Privacy Act (ECPA) provides a basis for seeking the information from the service provider. Typically, a litigant will send a letter of preservation to the service provider followed by a subpoena, court order, or search warrant.

One issue that comes up frequently in domestic relations cases is where the husband and wife have equal dominion over a computer: to what extent can one spouse spy on the other? There are a number of state and federal laws that, while not interfering with a spouse’s property rights concerning the computers or devices, do operate to protect the privacy rights of those who use the devices. For example, one man was convicted for unlawful interception of electronic communication under Texas law after his wife, with whom he was living and whom he was divorcing, revealed to police that he was intercepting her private phone conversations. Although the husband probably had “equal dominion” over the phone in his house, he was nevertheless sentenced to two years' imprisonment and a $1000 fine.

There is case law in California and Virginia clamping down on civil litigants’ rights to use the ECPA. In O’Grady v. Superior Court, the court held that civil litigants can no longer obtain information through civil subpoena and are forced to seek e-mails, etc., from senders.

85. 18 U.S.C. §§ 2510–22 (2012); see id. §§ 2701–12.
88. Id., at 19.
89. See O’Grady v. Superior Court, 44 Cal. Rptr. 3d 72 (Ct. App. 2006) (holding ECPA restricts access to ISP information to law enforcement only); In re Subpoena Ducas Técum to AOL, L.L.C., 550 F. Supp. 2d 606, 611–12 (E.D. Va. 2008).
90. Id.; see In re Subpoena Ducas Técum to AOL, L.L.C., 550 F. Supp. 2d at 611–12 (holding that receipt of civil subpoena does not authorize ISP’s disclosure of stored e-mails under ECPA). But see Romano v. Steelcase Inc., 907 N.Y.S.2d 650, 654–57 (Sup. Ct. 2010) (citation omitted) (ruling the SCA inapplicable because the information sought was both material and necessary to
D. Another Data Source—Call Detail Records (CDR)

The ECPA allows for the discovery of subscriber data and CDR from service providers. CDRs are produced in the telephone switch and can include the following types of information: date/time of call origination and termination, called and calling party, duration of the call, type of call (inbound, outbound), and the originating and terminating tower (base station). The information looks a lot like a phone bill.

CDRs are a potential form of evidence, but knowing the merits and limitations of CDRs is critical when using them in a case. CDRs need a significant amount of supplementary information to be of most use, such as the maintenance records or trouble tickets of the cell towers referenced in the call detail records, information about the configuration of cell towers of interest, and the radio frequency maps for the cell towers, if available. These and other factors must be taken into consideration and dealt with if an examination of CDRs is to be performed comprehensively and correctly. Certain items are not included in CDRs, including phone address book (contacts); calendar, tasks, and notes; photos

the defense of the action and could lead to admissible evidence, and that “[t]o permit a party claiming very substantial damages for loss of enjoyment of life to hide behind self-set privacy controls on a website, the primary purpose of which is to enable people to share information about how they lead their social lives, risks depriving the opposite party of access to material that may be relevant to ensuring a fair trial”).

93. See, e.g., United States v. Yeley-Davis, 632 F.3d 673, 678 (10th Cir. 2011) (finding admission of certified cell provider’s CDRs did not violate defendant’s rights under the Sixth Amendment’s Confrontation Clause); United States v. Sanchez, 586 F.3d 918, 928–29 (11th Cir. 2009) (affirming defendants’ criminal convictions and finding no abuse of discretion when the district court admitted cell phone company call records).
94. Terrence P. O’Connor, Provider Side Cell Phone Forensics, 3 SMALL SCALE DIGITAL FORENSICS J., June 2009, at 3.
95. LARRY DANIEL & LARS DANIEL, DIGITAL FORENSICS FOR LEGAL PROFESSIONALS: UNDERSTANDING DIGITAL EVIDENCE FROM THE WARRANT TO THE COURTROOM 163–64 (Robert Maxwell et al. eds., 2011).
and videos; audio clips and voice memos; browsing activity and apps; and anything deleted.

E. Text Messages

People are changing the way they communicate with each other. Over ninety percent of adults own mobile phones and eighty percent use their telephones to send and receive text messages. Many even prefer texting to talking. This trend is most pronounced among younger people who over time will comprise a greater percentage of the workforce. Ninety-seven percent of adults between the ages of eighteen and twenty-four have mobile phones and send or receive on average over 100 messages per day. Text messages are a potential source of relevant evidence and need to be accounted for in discovery.

Depending on the wireless company, an individual user usually can access his or her online account and print a list of numbers/dates/times text messages that were sent and received. Users may be able to obtain this information for some period of time. According to a 2011 Wired.com story, based on a leaked August 2010 Justice Department memorandum,

Verizon, for example, keeps a list of everyone you’ve exchanged text messages with for the past year. T-Mobile stores the same data up to five years. It’s 18 months for Sprint, and seven years for AT&T.

The biggest difference in retention surrounds so-called cell-site data. That is information detailing a phone’s movement history via its connections to mobile phone towers while its traveling.

Verizon keeps that data on a one-year rolling basis; T-Mobile for “a year or more;” Sprint up to two years, and AT&T indefinitely, from July 2008.

Unless saved on a user’s hand-held device, only the wireless service provider may have access to past text messages, and

96. Catherine A. Bernard, Text Messaging as Uncharted Territory in E-Discovery, CORP. COUNS. (Online) (Oct. 4, 2013), available at LEXIS.
97. Id.
providers are not required to maintain texts for any set period of time. There is a very short time span to recover the body of text messages from a provider. According to the ACLU of North Carolina, service providers generally keep text messages as follows:

- Verizon: 3 to 5 days;
- T-Mobile: 0 days;
- AT&T: 0 days;
- Sprint: 0 days;
- Virgin Mobile: 90 days.99

Given this short time span, in situations where litigation is reasonably anticipated, it is recommended that clients be advised in a litigation hold letter to preserve and not delete any potentially relevant text messages and that a request be sent to opposing counsel or an opposing party putting it on notice to preserve potentially relevant text messages. As text messages become a more prevalent form of communication, their value as potential evidence will continue to grow.

CONCLUSION

Although the evolution of technology has increased both the amount of digital information and sources where potentially relevant digital information is stored, technology is also evolving to assist in the more efficient assessment, review, and production of information. Meaningful discussions on these topics with clients and opponents must occur early in a case to most effectively ensure that electronic discovery costs stay proportionate to what is at stake in the litigation. With foresight and planning, costly and protracted disputes over document review and production can be avoided or minimized which might otherwise threaten the just, efficient, and cost-effective resolution of disputes.

APPENDIX A: PRODUCTION FORMATS

What follows is a list of the most common types of production formats, with a description for each.

Load Files

- A load file is a data file that defines the links between multiple records in a database and document images.
- A load file is used to import data about the documents and the images into a litigation database.
- Most common load files—Summation (DII, CSV) and Concordance (OPT, DAT).

TIFFs

- **TIFF (Tagged Image File Format):** The most widely used and supported graphic file format for litigation images. It is a pixel-by-pixel representation of a paper or electronic file—a “picture” of the document. It is good for black and white documents.
- **TIFF Group IV (compression):** A two-dimensional compression format for storing images. Typically compresses at a twenty to one ratio.
- **Single Page TIFF:** Each page of a document is a separate TIFF file (most commonly requested for litigation databases).
- **Multipage TIFF:** A single TIFF file containing all pages of a document.
- A TIFF image of a document cannot be altered, adding security, authenticity, and integrity to the document set for review and production.
- If native files are produced in image format, then hidden data, embedded text, tracked changes, or spreadsheet formulas may not be visible without access to native files.

PDFs and JPGs

- **PDF:** Adobe file format that allows a document to be viewed by any computer with the free Adobe reader regardless of what application was used to create the original document. PDFs can either be text based, if they were generated by the native
application as an alternative file format, or image based, if they originated as a scanned image of a hard copy document.

- **JPG/JPEG**: The standard image compression used for photographs and color documents. Requires less storage and quicker downloading than other photo file formats.

**Native Files**

- Each document is produced using the file format specific to the software used to create it—Word, Excel, PowerPoint, CAD, Outlook, etc.
- Documents can only be opened within the original application, but they can often be viewed with various viewing applications.
- Documents are “live,” meaning they are available for editing.
- All metadata is available—including track changes if the document was reproduced in a manner that captured all the pertinent metadata.

**TIFF/PDF vs. Native**

- **TIFF/PDF**
  - Easy to endorse with Bates numbers and confidentiality;
  - Uniform format;
  - Easy to redact;
  - Must pay to convert;
  - Can negotiate metadata fields to be produced;
  - Easily authenticated.

- **Native**
  - No conversion costs;
  - Live document can be altered—potential for spoliation;
  - Difficult to endorse each individual page with Bates numbers and confidentiality designations;
  - Carries all metadata—could be difficult to review if track changes was “on,” for instance.
APPENDIX B: HOW IS E-DISCOVERY DIFFERENT FROM DIGITAL FORENSICS?

<table>
<thead>
<tr>
<th></th>
<th>E-Discovery</th>
<th>Digital Forensics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relevant Community</strong></td>
<td>Legal—Civil Litigation</td>
<td>Law Enforcement</td>
</tr>
<tr>
<td></td>
<td>Records Management</td>
<td>Telecommunications</td>
</tr>
<tr>
<td><strong>Activity</strong></td>
<td>Collect, Review, and Produce</td>
<td>Acquire, Examine, and Report</td>
</tr>
<tr>
<td><strong>Custodians</strong></td>
<td>Many</td>
<td>Very Few</td>
</tr>
<tr>
<td><strong>Evidence</strong></td>
<td>Documents</td>
<td>Mobile—Phone, Tablet, GPS</td>
</tr>
<tr>
<td><strong>Formats</strong></td>
<td>E-mail and Attachments</td>
<td>Online—Web Mail, Social</td>
</tr>
<tr>
<td></td>
<td>PST, NSF, XML, TIFF, Native</td>
<td>XLS, CSV, HTML, PDF</td>
</tr>
<tr>
<td><strong>Tools</strong></td>
<td>Summation</td>
<td>UFED</td>
</tr>
<tr>
<td></td>
<td>Concordance</td>
<td>SecureView</td>
</tr>
<tr>
<td></td>
<td>Relativity</td>
<td>Device Seizure</td>
</tr>
<tr>
<td></td>
<td>Ringtail</td>
<td>Oxygen</td>
</tr>
<tr>
<td></td>
<td>Catalyst</td>
<td>MPE+</td>
</tr>
<tr>
<td></td>
<td>Clearwell</td>
<td>Nextpoint</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WebCase</td>
</tr>
</tbody>
</table>
APPENDIX C: SUBPOENA TO FACEBOOK

In a subpoena to Facebook, use the following technical language,
including the identifying information for the user’s profile:

- For the Facebook user account identified by the Facebook ID https://www.facebook.com/user.name, birth date of October 12, 1963, with the following e-mail addresses that may be connected to the Facebook user account:
  - email@myemailaddress.com;
  - mymail@somefreeemail.com; or
  - email.address@someotheremail.com.

- For the period of January 1, 2009 through May 1, 2010:
  - All activity for the user account, including wall posts, chat logs, profile and album pictures, friend lists, and profile pages.
  - Original creation date of the user account and profile.
  - A log of all IP addresses used to access the account with the date and time for each access, including the MAC address of the connecting computer for each connection.

100. See DANIEL & DANIEL, supra note 95, at 140–41.
101. Id. at 141.
102. Id.