

Efficiency: A Discovery Philosophy, and All You Really Need to Know About Predictive Coding

December 19, 2016

The main problem with discovery is the cost. In a very small number of truly bet-the-company cases (for example, where the CEO's emails must be produced) the greater risk can be failing to do discovery perfectly. But 99 times out of 100, cost is the most important factor in discovery. My guiding principle in handling discovery is, therefore, to reduce cost. One of the biggest problems with the way many lawyers approach discovery is that they perform its steps in the wrong order. The order of steps in discovery can have an enormous impact on its costs. Meaning, many law firms will get a case in the door, then begin thinking about discovery, then send and receive discovery requests, and only then approach the client to begin to determine what documents the company actually has, and what systems and capabilities it has to retrieve them. The lawyer then begins a good-faith conference process to negotiate with opposing counsel what procedures will be used in discovery, only after objections have already been sent. This means the lawyer will often have written responses to discovery requests, and objections to those requests, without knowing what the company actually has or what problems it may face in retrieving documents from its systems. In today's world this is simply irresponsible. The better order of steps is to: receive the case, do a preliminary analysis of what will likely be relevant and what the other side will likely want. Then, contact the company and discuss this preliminary analysis, begin to determine what documents the company has, and gain an understanding of the company's systems. The lawyer should begin the good-faith conference process with opposing counsel before any discovery requests are sent or received. This means negotiations with opposing counsel will let the lawyer and the company know how reasonable the other side intends to be (and reveal the likelihood that discovery will become a problem) before discovery responses are even prepared. If the other side plans to be obstreperous, then it becomes important to frame discovery responses and objections carefully so they will look good to the court in motion practice. However, if the other side plans to cooperate, a more cooperative spirit governs responses. The tone is different tone, and far less time and money can be

spent on the discovery responses. Only after these factors are known, and the company's documents and systems have been taken into account, should discovery responses be sent or responded to. Another difficult aspect of discovery for many companies is uncertainty associated with what it will cost to respond to requests, gather responsive documents, get them produced, and, if necessary, persuade the court that the company has fulfilled all its discovery obligations. The best way to deal with uncertainty is to use your experience in discovery to establish either capped fees or piece-rates that are fair, reasonable, and give the company a measure of certainty in advance about discovery costs. I believe the best way to establish such rates is transparently and collaboratively with the company. For some large-scale discovery projects, the way to save money is to use a document review vendor, or computer assisted review. For matters in which the company expects to review hundreds of thousands of document it should be company policy to employ outside legal consultants to conduct document review, rather than having higher priced outside counsel be solely responsible for all aspects of review. Some outside counsel resist this idea because they don't want to give up control. But with reviewers offshore charging \$50 per hour, the efficiencies can be tremendous, and outside counsel should overcome their fear. In the largest of large scale projects, computer assisted review tools are likely to become more efficient than even a document production consultant. In some matters, it is appropriate to use both computer-assisted review tools and a document production consultant. But in many matters using computer-assisted review tools can obviate the need for a document production consultant. The following discussion of computer-assisted review tools will familiarize you with what they are and what they do.

Understanding computer assisted review Certain important standard technical terms and techniques have been developed in the electronic discovery literature. See Aaron Goodman, "Predictive Coding," 43 *Litigation* 23 (Fall 2016); Lea Malani Bay et al., "Technology-Assisted Review: Advice For Requesting Parties," *Practical Law* (Nov. 2016). In my opinion, the best way to explain these terms and techniques to the court is to submit an affidavit that defines and explains them using common sense examples. You can save money, if appropriate, by having a company employee submit this affidavit. But if litigation is more contentious or involves higher stakes, it might be worthwhile to have an independent electronic discovery expert submit this affidavit. Here are the most important terms and techniques: **Proportionality**: the legal concept that discovery should not outweigh and overwhelm the value of what is at stake in litigation. Proportionality is now specifically mentioned in the Federal Rules of Civil Procedure. The concept is also present in the reasoning of many state court decisions, although most states have not yet adopted their Rules of Civil Procedure to mirror the federal rules. **Good faith conference**: the procedure by which parties meet, discuss the discovery needs, and try to agree on particular methods that will be used to identify relevant documents for production. The Federal Rules of Civil Procedure require good faith conference, and spell out much of what parties are required to do with respect to conferring about electronic discovery. In addition, a growing number of state courts also require good faith conference to one extent or another. While some litigants take an aggressive or obstreperous approach to good faith conference, my opinion is that aggressiveness only costs more money and rarely advances the company's interests. A transparent and cooperative approach to good faith conference can save tremendous amounts of

money, and even if agreement cannot be reached, puts you in the best position to win discovery motion practice before most courts. It is rare that a transparent and cooperative approach will reveal attorney work product or compromise the company's approach to litigation, which is often the concern litigants have with this approach. So, for example, sharing the list of custodians, sharing a list of repositories, trying to come to a reasonable agreement on keywords, and being reasonable with the use of technology assisted review for both sides will, in my opinion, benefit the company by saving money and making victory more likely. *See Hyles v. New York City*, 2016 WL 4077114 (S.D.N.Y. Aug. 1, 2016) (an example of contentious discovery driving up costs that involves electronic discovery and technology assisted review); *United States v. Education Management LLC*, 2013 WL 12140442 (W.D. Pa. Nov. 24, 2013) (same); *see also Apple, Inc. v. Samsung Electronics Co. Ltd.*, 2013 WL 1942163 (N.D. Cal. May 9, 2013) (explaining parties' duty to confer cooperatively and transparently about electronic discovery); *Romero v. Allstate Ins. Co.*, 271 F.R.D. 96 (E.D. Pa. 2010) (same).

Clawback: an agreement by the parties, sometimes enforced by a court order, that if privileged documents or work product are inadvertently produced, the party receiving the privileged documents will return them and not argue that the privilege has been waived.

Custodians: all of the people within the company who may have relevant documents.

Repositories: all the places (email boxes, files, folders, etc.) where custodians may keep potentially relevant documents.

Universe: all the documents that you intend to review. In other words, your universe contains all of the documents found in all of the repositories belonging to all of the custodians.

Keyword search: a review of the universe that searches for particular words and returns only those documents that contain those words. The usual criticism of keyword searching is that the people trying to think of the keywords will not be able to come up with all of the terms that may be found in relevant documents, so relevant documents can be missed. Another weakness of keyword searching is that some keyword search tools can search only a precise word. So, for example, if you search for "closings", but the word "closing" appears in your document, the search tool may miss it.

Boolean search: a type of search that addresses this last problem. More sophisticated systems are able to use things like connectors, wildcards, etc. that ensure the computer will return documents similar to the keywords you used, even if not an exact match. This can dramatically improve the effectiveness of a keyword search.

Technology assisted review: sometimes called TAR, this is any tool used for a review in which a computer is trained to identify relevant documents, so your review team need not manually review the entire universe.

Predictive coding: the most common and well-accepted form of TAR. In predictive coding, a review team reviews a set sample size of documents from within the universe, their coding is fed into the computer, and the computer reviews the rest of the universe and returns relevant results.

Seed set: a group of documents, sampled from the universe, that is manually reviewed to determine relevance, then coded and used to train the predictive coding software. The seed set is sometimes reviewed only by the producing party, but sometimes both producing and requesting parties can agree to cooperate and review the seed set together. This can minimize the likelihood of discovery disputes and save cost.

Computer assisted learning: a less common and less well-accepted form of TAR, but one that may represent the future because the empirical literature suggests it may be more effective. Sometimes called CAL, it also involves training a computer to

identify relevant documents, but it does not begin with a fixed seed set. Instead, it continues to actively train as reviewers do their work. **Recall:** the percentage of relevant documents in the universe that are identified in your review. For example, if you have a universe of 100,000 documents, and 30,000 of those documents are actually relevant, and your review identifies 15,000 documents as relevant, then your recall is 50 percent. **Precision:** the percentage of documents your review identifies as relevant that are in fact relevant. So if your review identifies 30,000 documents as relevant but a quality control review shows that only 10,000 documents are relevant then your precision is 33 percent. **Richness:** the percentage of documents in your universe that are actually relevant. So if you have a universe of 100,000 documents and 30,000 of those documents are actually relevant, then your richness is 30 percent. **F1:** a highly technical term that can be difficult for courts to understand, but it is the gold standard for measuring the quality of a computer assisted review overall. It is the “harmonic mean” (a particular kind of average) of your recall and your precision. It is intended to measure effectiveness in a practical and conservative way, so it tends to be closer to the lower (less effective) of your recall and your precision. **Nested review:** a system in which manual review, keyword searching, and technology assisted review, are used together in an agreed-upon order to identify all the relevant documents within the universe that will be produced. For example, a review might begin by running a Boolean keyword search on the email boxes of all of the document custodians to generate a subset of documents. That subset will then be reviewed using predictive coding to winnow the documents further. All of the documents identified by these two techniques as relevant might then be reviewed manually to identify privileged documents and make final relevance calls before a production is made. See *In re Lithium Ion Batteries Antitrust Litig.*, 2015 WL 833681 (N.D. Cal. Feb. 24, 2015) (describing a typical nested review); *Progressive Cas. Ins. Co. v. Delaney*, 2014 WL 3563467 (D. Nev. July 18, 2014) (same). A large and growing body of empirical research shows that technology assisted review is actually more effective than manual review of the entire universe of documents. See Maura R. Grossman et al., *Technology-Assisted Review In E-Discovery Can Be More Effective And More Efficient Than Exhaustive Manual Review*, 17 Rich. J. of Law & Tech. 11 (Spring 2011); Nicholas Barry, “Man Versus Machine Review: The Showdown Between Hordes of Discovery Lawyers and a Computer-Utilizing Predictive-Coding Technology,” 15 Vand. J. Ent. & Tech. L. 343 (Winter 2013). In other words, computer tools have gotten so good that they can produce better recall, better precision, and better F1 than a team of manual reviewers who actually lay eyes on every document in the universe. This is an amazing feat. It is the reason why courts today generally accept technology assisted review when the parties in litigation agree to its use. See *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012) (the seminal opinion by widely respected Sedona-conference active Magistrate Andrew J. Peck that first approved predictive coding in court); see also *Rio Tinto PLS v. Vale S.A.*, 306 F.R.D. 125 (S.D.N.Y. 2015) (Magistrate Peck’s follow-up to *Da Silva Moore*). Although it is much less common for parties or courts to consider using technology in assisted review if either side insists on manual review, this is probably the future. See *Hinterberger v. Catholic Health System, Inc.*, 2013 WL 2250603 (W.D.N.Y. May 21, 2013) (refusing to compel the use of TAR over the other party’s objection); *Bridgestone Americas, Inc. v. IBM*, 2014 WL 4923014 (M.D. Tenn. July 22, 2014) (ordering parties to confer in good

faith about TAR); *but see also Dynamo Holdings Ltd. Partnership v. Comm. of IRS*, 2016 WL 4204067 (U.S. Tax. Ct. July 13, 2016) (after parties agreed to TAR, one party could not compel the other to do a further review). **Conclusion** In an appropriate case we would urge a court to order the use of technology assisted review, even against opposition, to save money, make litigation more efficient, and be more consistent with the proportionality concept. I believe that if all these tools, techniques, and services are used, we can maximize the likelihood that the company will achieve optimum efficiency in driving down the total cost of its discovery throughout the nation.

Related Practices

[Real Estate](#)

Related Industries

[Real Estate](#)

©2024 Carlton Fields, P.A. Carlton Fields practices law in California through Carlton Fields, LLP. Carlton Fields publications should not be construed as legal advice on any specific facts or circumstances. The contents are intended for general information and educational purposes only, and should not be relied on as if it were advice about a particular fact situation. The distribution of this publication is not intended to create, and receipt of it does not constitute, an attorney-client relationship with Carlton Fields. This publication may not be quoted or referred to in any other publication or proceeding without the prior written consent of the firm, to be given or withheld at our discretion. To request reprint permission for any of our publications, please use our Contact Us form via the link below. The views set forth herein are the personal views of the author and do not necessarily reflect those of the firm. This site may contain hypertext links to information created and maintained by other entities. Carlton Fields does not control or guarantee the accuracy or completeness of this outside information, nor is the inclusion of a link to be intended as an endorsement of those outside sites.