

# Controlling E-Discovery Costs in a Big Data World

by Brian Ingram of LexisNexis

## *Legal IT professionals are all too aware*

that the task of managing large mountains of data has always been a defining characteristic of our industry. Now that we've had to layer on the explosion of digital and social media content, we've got a whole new set of challenges in front of us. The attorneys we support are looking to us for new approaches to help them make sense of all this data.

## **The World of Big Data**

In particular, lawyers are taking greater notice of the legal, ethical and technological issues around the new frontier for information technology: big data. IT professionals are on the leading edge of this new battleground, developing innovative approaches to information governance and implementing sophisticated technology systems to harness the power of this data avalanche.

One promising application for big data analysis is in electronic discovery, where fast, high-performing data analytics can substantially reduce the time and cost of preparing for a case. In fact, at the root of any e-discovery project or process is the ability to identify, collect, index and analyze big data. As many legal IT professionals are discovering, new analytics tools and big data technology innovations are making e-discovery software smarter, which can only help law firms and their clients better manage

risks associated with failing to produce all relevant documents in a lawsuit or investigation.

However, big data not only bring improved tools to manage large quantities of information, there are also new challenges to identify the most relevant data in e-discovery. The process involves gathering information from a diverse range of sources, storing that content and then developing systems so the information can be mined, analyzed and produced. Big data require an IT environment that can easily scale to handle high electronic data discovery processing without experiencing technical problems.

For law firms of all sizes, there are four key "best practices" for conducting e-discovery more efficiently while controlling costs in the big data environment.

### **1. Develop a Strategy for Information Governance**

It's important to start out by creating and putting into place a comprehensive data management program for compliance with regulations, statutes and best practices. This should involve the development of customized guidelines and procedures for the creation, storage and disposition of any and all types of data. Also create email policies, litigation hold procedures and disaster recovery plans.

Many organizations find that one way to improve efficiency and reduce costs is by performing an inventory of enterprise data

— developing a data classification process and creating a data retention policy. You can then develop organizational management policies and procedures for electronically stored information (ESI) that include electronic email policies and develop workflows to deal with the potential for large amounts of nonsearchable data, including hard copy documents.

Finally, turn your attention to security issues as part of your information governance strategy. Stay current on regulatory and legal data security obligations so you can create a data security approach based on repeatable and defensible best practices. Be sure to develop a data breach response plan that is unique to your organization.

When the information governance strategy is in place, it should look something like this:



### 2. Establish Rules for Data Extraction and Collection

While there are numerous methods used for data collection in e-discovery, they can be narrowed down to two general categories: automated and manual. With the explosion of big data, we have seen a proliferation of automated ESI collection tools on the market. While these tools have helped manage an enterprise's data and provided more streamlined methods for the collection of ESI, many firms still opt to go the manual custodian, "self-collection," route.

The courts have yet to stake out a clear position on either method of collection, but the Sedona Conference has provided some general guidance in this area. In their "Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery," practice point one states: "In many settings involving electronically stored information, reliance solely on a manual search process for the purpose of finding responsive documents may be infeasible or unwarranted. In such cases, the use of automated search methods should be viewed as reasonable, valuable, and even necessary."

Regardless of the collection methods used, it is the education about, compliance to and ongoing review of your data management policies that are the critical tasks that should be undertaken by corporations and outside counsel to maintain a repeatable and defensible ESI collection process for e-discovery.

---

Develop workflows to deal with the potential for large amounts of nonsearchable data, including hard copy documents.

---

### 3. Prioritize Data Sets

There are three primary methods used to prioritize and organize data sets to be reviewed by counsel:

- Keyword searching
- Concept searching
- Use of analytics or predictive coding technology

Each approach has its own unique set of advantages and disadvantages and, when used in conjunction with careful documentation and appropriate iteration, should survive a challenge by a party's opponent in litigation. The optimal workflow will likely require a hybrid approach applying two, or perhaps all three, of these approaches to reach the best results.

For example, the use of analytics and predictive coding typically returns more relevant documents without the limitations of traditional searching logic. However, a well-planned and well-constructed keyword search can be more effective for certain situations, such as locating specific date ranges and/or proper names. Putting these technologies and methods together with the proper workflow, methodology and documentation will deliver the best results and better maintain defensibility if challenged.

More important than the actual technology is how the legal teams — and the professionals who are familiar with system limitations — combine them into the optimal workflow best suited to meet the requirements for a given case. In the end, the success of these approaches is based more on the personnel and workflow and not the technology itself.

### 4. Select the Best Technology Tools

Many law firms and organizations have responded to the explosion of data and the resulting increase in litigation matters by choosing to invest in their own e-discovery software tools to manage tasks. With so many software options available, companies have to evaluate their needs carefully before determining the right e-discovery software for their specific requirements.

## Controlling E-Discovery Costs in a Big Data World

Every organization is unique, and there isn't a "one size fits all" option available. Therefore, the evaluation process of e-discovery solutions that are affordable and guarantee a good return on investment is of the utmost importance.


The first step in evaluating the actual tools is to decide which part of the e-discovery process you want to invest in. Some firms have decided that having ESI processing and analysis capabilities in-house is their priority over developing their own robust document review platform. Others have determined they are best served by outsourcing their processing needs and prefer to host and manage their own document reviews internally. Still others have made the decision to invest in all areas of the e-discovery continuum and take greater ownership of the entire process. Only the professionals inside each individual organization are capable of making these decisions, and they should only be made after careful and thorough analysis.

Any investment in e-discovery technology can be quite expensive, and you don't want to lock yourself into a solution that won't deliver what you need. These systems should be fast, powerful and versatile enough to process and manage a large and diverse population of data. Surveys have shown that lawyers' top priorities are speed and ease of use when using document review software. Selecting a tool that contains the features and functionality your lawyers and review teams desire while maintaining the necessary speed and ease of use will go a long way toward ensuring a successful implementation.

Scalability and flexibility are other important factors in selecting the right tool. A typical challenge for many organizations handling large projects is ensuring their tools can scale to handle very large volumes of data in a short time frame. In addition, a tool that has the flexibility to be customized to meet the unique and specific needs of any matter can prove to be very beneficial.

### More Data Breeds More Efficiency

Litigation discovery has always been an intensive process. The traditional approach was largely a waiting game of putting off discovery until it was finally required to begin the time-consuming process, oftentimes with attorneys hunched over boxes of documents for hours at a time. Legal IT professionals can attest better than anyone, the accelerating volume of data creation in the digital age has simply outpaced our ability to manage it effectively during litigation discovery, which is now more complex and costly than ever.

The good news is that the brave new world of big data is leading us to establish clearer rules for data extraction, more logically prioritized data sets, smarter information governance strategies and better technology tools. This may not tame the explosion of electronic data in the litigation world, but it does provide hope for navigating those mountains of data better while managing risk and making the e-discovery process more efficient. 



*Brian Ingram is a litigation technology and e-discovery professional with over 20 years of experience in the legal profession. Having spent the majority of his career managing litigation support departments at large law firms, Brian has been involved in some of the largest matters in the nation involving electronic evidence. He has also spent time on the corporate side of the industry. Brian is currently head of the Litigation Technology Consulting division at LexisNexis. He can be contacted at [brian.ingram@lexisnexis.com](mailto:brian.ingram@lexisnexis.com).*



This article was first published in ILTA's March 2013 issue of *Peer to Peer* titled "Data Management: Taming the Explosion" and is reprinted here with permission. For more information about ILTA, visit their website at [www.iltanet.org](http://www.iltanet.org).