An Introduction to Electronic Discovery for the German Legal Market

An AccessData/reThinkLegal White Paper

American lawyers have found a useful strategy that is gaining increasing acceptance in Germany: the use of electronic discovery (or "e-discovery") software in order to deal with huge amounts of data and to extract relevant information.

"Discovery" is part of the preparation of civil litigation in the U.S. Each litigant is obligated by law to provide certain data to the other party, which means that hundreds of thousands of documents are often exchanged. And since most employees now communicate via email, social media and messenger services, it's really the "electronic" in discovery that becomes relevant.

E-discovery software collects, extracts, and organizes the data from all kinds of servers and devices. The attorney may use e-discovery to redact PII (Personally Identifiable Information), company secrets and the privileged communication between employees and the company's attorney before handing over the data to the other party. Furthermore, the attorney may review the documents and tag them as "relevant" or "not relevant" in order to proceed with the relevant documents for his own work.

> **No matter which tools an organization uses, at a minimum the data must be identified, collected, processed for analysis, internally reviewed and culled, and then processed again.**

Even though the obligation to provide certain data to the other party before the actual litigation is unknown in Germany and other parts of the EU, e-discovery has vast potential to support the work of German attorneys simply because e-discovery software is capable of dealing with large amounts of data. These tools organize data and make it easy to access, not only for litigation matters, but also for antitrust audits, compliance reviews and forensic investigations.

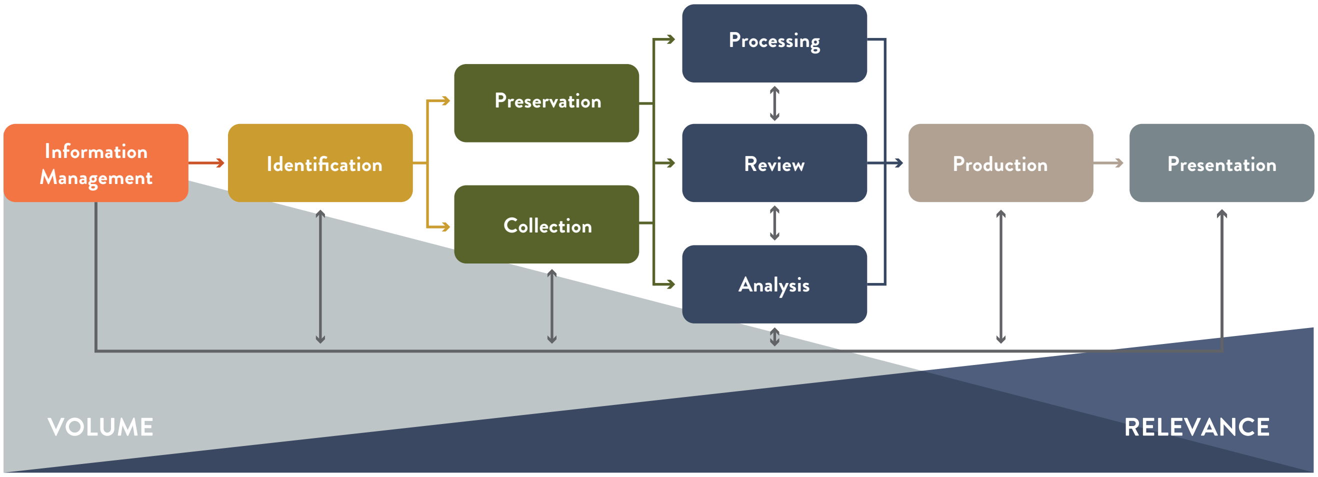
The E-Discovery Workflow

Traditionally, no matter which tools an organization uses to address electronic discovery, at a minimum the data must be identified, collected, processed for analysis, internally reviewed and culled, and then that reduced data set must be processed again for import into to a legal review tool for outside counsel.

Some organizations use three or four different products, one to address each phase. In this case, the data must be processed each time it moves from one tool to the next—so any analysis, comments and tagging done during the internal review phase are lost when the data is imported into the legal review product. But many corporations have discovered the benefits of cost-efficiency and risk management that can be realized by using a single software platform to manage the entire spectrum of evidence collection, processing, review and production.

A non-profit organization in the U.S. was formed to study and create a model that captures this entire workflow, known as the Electronic Discovery Reference Model (EDRM). Here is an illustration of that model:

Electronic Discovery Reference Model



Before even entering the realm of EDRM preparedness, an organization should consider its system management and security needs. At a minimum, any software the firm purchases should accommodate flexible user access, including the ability to support the existing roles and responsibilities of the organization.

Consideration should be given to:

- User roles and access rights should be customizable and not force the organization to adopt a specific workflow or team makeup.
- The solution must provide a rich set of system audit and logging reports to determine user activity at a specific date and time. Access to the login feature should be severely restricted to serve as valid input for establishing chain of custody, as well as supporting user management activities by showing who has what access.
- The system's communication and security protocol must also be robust and support the organization's needs and current configuration. The security feature's overarching function should be to prevent unauthorized use of the system or system components.
- Workflow steps should be clearly broken out and allow for tasks to be input into the system without an undue burden on multiple parties. Further customization to fields within the user interface should support organization requirements and changes.

Identification

The EDRM "Identification" phase includes development of a plan, as well as determination of sources for potentially relevant electronic evidence.

One of the essential parts of the identification phase is custodian tracking, because it helps organizations be proactive about where information resides and facilitates litigation preparedness. Historical custodian detail should be readily accessible to understand what other cases/matters a custodian was involved in and what data was collected in conjunction with that matter. Commonly stored data will include custodian email, phone numbers, business unit, case and other pertinent workflow details, which the system should be able to preserve and organize. Part of the system's ability to track custodians should be supported by structured data connectors, which should integrate between the system and the firm's existing programs.

The Data Map may be the most crucial part of the Identification phase because knowing where the data is and its accessibility level is intrinsic to planning the entire case strategy. Thus the data mapping functionality should be robust and should include the ability to record and update potentially relevant data repositories, such as PCs, email systems, SharePoint®, archiving systems and other structured data repositories. It is essential that this functionality be built in so that organizations do not have to incur the additional expense or complexity of using a third-party provider. Hand-in-hand with the Data Map is the pre-collection audit capability, which allows organizations to survey their information universe before they start the onerous work of collecting.

Support for this functionality should include the capability to run search criteria against a potentially relevant target and provide results without actually copying the underlying files. The pre-collection audit option is preferable to solutions that have to pull data back to index and report on potential search criteria because it is much faster. This can be a key advantage especially early in a case or when tight deadlines are approaching. Pre-collection auditing minimizes system and network impact and eliminates overhead because it reduces the storage of extraneous data for early case assessment activities. The importance of the pre-collection auditing capability should not be understated because the practice will continue to grow in significance as organization data grows in size and complexity.

Preservation

The EDRM "Preservation" phase includes data isolation and notification to appropriate parties that data related to an upcoming law suit must be preserved.

The largest, most important part of the preservation functionality is the Litigation Hold System. This piece of the software comes into play the moment litigation is contemplated by the parties and thus should lay the foundation for a highly organized and efficient case workflow. Litigation hold functionality should be fully integrated and not require third-party add-ons. This is because integration provides the benefit of centralized management of all custodians and eliminates the complexity and cost of introducing a distinct system into the IT environment.

Some essential components of a comprehensive litigation hold system include:

- Up-to-the-minute progress tracking of all statuses within the matter (including custodians and IT specific managers or data owners)
- Optional workflow approval sequence for attorney or paralegal review
- Templates and other customization tools to increase efficiency
- Distribution and management of attachments and custodian interview Q&A recording and support
- The system should be able to produce electronic and hard copy reports for use internally or with external parties

Collection

The EDRM "Collection" phase includes acquisition of potentially relevant electronically stored information (ESI). Collection should include the document/file as well as any associated metadata.

Collection is a crucial part of the e-discovery process, which is reflected in the wide spectrum of offerings and definitions in this area. Many providers offer some level of collection, but few have years of experience and a solid track record of delivering defensible results. Collection and Processing capabilities should be heavily scrutinized to separate inflated marketing spin from the real thing. Organizations should take particular care to test and ensure data is not being dropped or missed (open files and email, system files, large files, etc.) during collection.

While certainly not a requirement, forensic data collection inherently achieves a degree of defensibility not available in a non-forensic collection. Forensic collection has other advantages as well as heightened defensibility, including the ability to audit the collection and the ability to collect deleted files. No longer solely the domain of law enforcement, forensic collection is rapidly becoming understood and sought by opposing counsel and the courts. The organization that chooses a tool with forensic collection capability not only chooses the strongest level of collection stability, but puts itself at the front of a developing trend. Whether the chosen solution offers a forensic collection capability or not, the collection solution must have a certain set of functionality in order to be minimally acceptable. Organizations should review and ensure their software tool has the ability to collect open files or files currently in use. Tools that fail to meet this critical criteria fall short of being legally defensible and leave organizations open to charges of incomplete preservation.

Also important, since most organizations have many potential custodians located offsite and outside the corporate network, is the ability to collect from employee laptops that are not logged in to the corporate network.

Processing

The EDRM "Processing" phase typically includes indexing, itemization and some level of data identification within the subject data universe.

The processing phase is the real workhorse of the e-discovery lifecycle. Within this phase, all data that was collected previously gets extracted and turned into information that can be culled down for greater relevance and read by review platforms in the next phase. As such, speed and accuracy are at a premium and a great deal of marketing dollars has been spent on claims related to data processing speeds (e.g., TB/day). The reality is that most of these claims are made using state of the art hardware platforms (prior to any licensing fees). The ideal software solution should be one that can easily and affordably scale using existing hardware to achieve processing speeds of terabytes per day.

As always, the best advice is to run a thorough POC (proof of concept) with your own exemplary data set and a full understanding of the service level objectives within your company. The single criterion that carries the most variance across e-discovery vendors is processing diligence (accuracy), meaning the thoroughness and accuracy of the processing tool. Because processing happens "under the hood," data can easily go missed and undetected or unreported to end users.

Other key capabilities for your e-discovery software solution include:

- Integrated Optical Character Recognition (OCR); the ability to extract text from document images or PDFs so that it can be searched in subsequent e-discovery phases.

- The program should be able to perform full text extraction from electronic documents and email to facilitate the same.

- The application should support a full range of document deduplication (identification of exact duplicates) options and should flag and optionally remove duplicates, then generate reports showing which documents/emails are duplicates with associated counts.

Analysis

The EDRM "Analysis" phase includes evaluating the collected and processed data to determine overarching information about key case topics, players and documents. For the purposes of this guide "Analysis" is synonymous with Early Case Assessment and Early Data Assessment.

The Analysis or Early Case Assessment ("ECA") phase of the e-discovery process entails taking the large and unorganized set of data from the processing phase to determine what type of case you have and whether you should go forward with the discovery process or look at settling. For this reason, analysis tools support the functions of categorizing, refining and bucketing data. The most well-known function is keyword searches and culling. All software applications and support processes should have an efficient and effective method for using keywords to analyze and reduce the subject corpus of data down to a manageable subset.

In addition, your software and workflow should include the following: in-document hit highlighting, keyword counts and summaries; data and evidence bookmarking at the global and case level to support categorization and organization; predefined "buckets" and document categorization, which allow the user to apply broad filters, such as file type or a date range to the data set; threaded view of email is also intrinsic to a quick and holistic analysis of the data; and comprehensive support for most legal review tools.

Finally, analysis and reporting in the ECA/EDA phase should be able to quantify and present which documents did and did not meet search criteria. These reports and metrics are critical input to further development of overall case strategy and can influence whether or not the user performs additional collections. Reporting can also help to quickly determine if chronological or conceptual gaps exist in the current data set.

Review

The EDRM "Review" phase focuses on sub-categorizing documents to identify relevant facts, further refine case strategy, and reduce risk to the client. Review is generally conducted by an attorney or other skilled practitioner.

The EDRM Review phase is the stage where documents receive the most scrutiny by the most highly trained (legally, not technically) users. Therefore, clarity and ease of use are extremely important, as these practitioners need to spend more time on legal analysis and less time on mastering software. To be ideal, the technology system you deploy must give a very detailed picture into the evidence and individual documents, yet be extremely intuitive.

Many successful systems employ a multiple-tier user interface that can accommodate both novice users and those needing advanced functionality. This helps to solve the business problem of requiring clarity and ease of use with the ability to support all case needs completely within a single tool. The multi-tiered user offering should be administered via granular security permissions and grouping structures that allow case managers to be flexible in creating review hierarchies and strategies. These permissions should be easily configured at any point in the review cycle and should include the ability to restrict or allow access to all software functionality.

Some of the basic components review software should offer include the following:

- Large and clear document viewer that can be undocked to move to a split screen
- Flexible review screen allowing the user to design an optimal viewing area with document summary
- Image and tagging view; a near-native document viewer that allows the user to view multiple (preferably hundreds of) file types without requiring installation of the native application, thereby speeding review time and significantly reducing cost and installation complexity
- Bulk tagging/coding of document groups and document families

It's also important for your technology tools to provide Unicode support to enable viewing and searching of foreign languages. Your litigation support team should be prepared to use redaction and document marking support, including text overlay, as well as custom color schemes. Redactions should appear transparent when performing review but be optionally "burned-in" at time of production so your team can review the underlying text without disclosing it at production.

Production

The EDRM "Production" phase encompasses export and exchange of electronically stored information in response to a production request between parties.

Production completes the arc of the e-discovery process, but includes much more than just printing documents out of a review platform and attaching a privilege log. Today's productions come in many formats and some may never see paper. Therefore, an organization's production capability has to be capable of handling not only the traditional production duties of redacting, printing and numbering, but be able to produce data in its many formats and load file iterations. The production workflow should also be flexible and allow for the creation of empty production set "buckets" to which documents can be added or from which documents can be removed as case objectives change.

While production is not quite synonymous with export, the way data gets out of a system is still an intrinsic part of the production process. Data should be available to export from the system in various formats, including load files, native files, images or forensic containers (supports portability or in cases of criminal matters). Exporting data should not incur an additional expense/fee or require the use of a third-party application.

Since the European legal world is far from giving up paper, production features for the classic method should be strong. An application should include the ability to burn in redactions, mark-ups and stamps and to Bates number productions sequentially. The stamping/Bates number functionality should include header or footer placement, a prefix option and a starting sequence and padding. The option to start numbering from a previous production set must be included for proper data management.

Along with supporting the output of documents from the system, a good product will also give the user ways to manage and track productions throughout the life of the case. For example, load file volume and document options support should be included for any number of foldering options to support work with an outside vendor upstream or downstream of the production phase. Also, custom data columns should be present in the case database to assist in meeting the needs of existing processing systems, outside counsel and partners.

Conclusion

The use of e-discovery software is common among American lawyers as it's the only realistic way to comply with American litigation rules when trying to deal with huge amounts of data and to extract relevant information. However, this same technology has huge potential in Germany and other EU countries as a way to make it easier to organize and access data involved in key legal functions such as antitrust reviews, compliance audits and forensic investigations.

ACCESSDATA® RETHINKLEGAL

AccessData Group has pioneered digital forensics and e-discovery software development for more than 25 years. Over that time, the company has grown to provide both stand-alone and enterprise-class solutions that can synergistically work together to enable both criminal and civil e-discovery of any kind, including digital investigations, computer forensics, legal review, compliance, auditing and information assurance. More than 130,000 customers in law enforcement, government agencies, corporations and law firms around the world rely on AccessData® software solutions, and its premier digital investigations products and services. AccessData Group is also a leading provider of digital forensics training and certification, with its AccessData Certified Examiner® (ACE®) and Mobile Phone Examiner Certification AME programs. For more information, please go to www.accessdata.com.

The team behind reThinkLegal has been working together since 2005. Their mission is the optimization of legal working processes. With the newest IT applications, legal expertise and professional project management, reThinkLegal is a full service provider to legal and compliance departments, law firms and agencies.Their way of operating is not only characterized by the team's legal background, but also by their economic understanding. For more information, please go to www.rethinklegal.de.

©2015 AccessData Group, Inc. All Rights Reserved. AccessData, AccessData Certified Examiner and ACE are registered trademarks owned by AccessData in the United States and other jurisdictions and may not be used without prior written permission. All other marks and brands may be claimed as property of their respective owners. 122015

Global Headquarters
+1 801 377 5410
588 West 300 South
Lindon, Utah

North American Sales
+1 800 574 5199
Fax: +1 801 765 4370
sales@accessdata.com

International Sales
+44 20 7010 7800
internationalsales@accessdata.com

LEARN MORE
www.AccessData.com