SPECIAL
REPORT

Bringing E-Discovery In-House:
RISKS AND REWARDS

By George J. Socha, Jr., Esq.

Socha Consulting LLC
www.sochaconsulting.com

Sponsored by:

ZyLAB®
eDiscovery & Information Management

Copyright

Table of Contents

Management Summary

This report provides a straightforward, pragmatic overview about how legal professionals and organizations confronted with e-discovery must be able to interpret e-discovery within the context of actual expected processes, inherent risks, and the available technical solutions that can support relevant activities. Many people may have some idea about what e-discovery is, at least thematically, but many do not have a full appreciation of how to effectively engage the setup and execution of the process. Even for those who have gone through an e-discovery process in the past, some of the acknowledged approaches to e-discovery are outdated, particularly when viewed against the current economic backdrop and the rapidly expanding technical challenges found in most organizations and legal firms.

Simply put, given the litigious nature of the marketplace and the disclosure and transparency responsibilities facing most organizations, no one at the front lines of their business can afford not to be well-versed about what e-discovery actually is and how it is actually conducted in a legal environment. (A more detailed situational overview is provided in the Chapter 2 of this document.) Moreover, the days of simply being able to hand off all of your e-discovery "problems" to a third-party has become less feasible due to exorbitant costs and increased legal risk associated with relinquishing control of information and discovery activities.

As such, organizations are desperate to find ways in which they can regain control of their internal processes in general – and e-discovery in particular – and minimize their costs and level of risk. A first step for organizations is to gain clarity about the actual issues that are in inherent in e-discovery. Chapter 3 of this paper provides an in-depth look at the key cost and risk-related issues that can compromise the effective execution of e-discovery processes.

After spelling out the basic risk and cost components of e-discovery, Chapter 4 follows with some best-practice approaches to remedying the most relevant e-discovery-related challenges. First and foremost, any effective e-discovery solution must have some reference point from which to understand the expectations and recognized process components that need to be engaged. This reference point is the Electronic Discovery Reference Model (EDRM), which is widely acknowledged as the standard by which e-discovery activities should be structured. Any technology solution used to support e-discovery activities must align itself as much as possible with the components found in the EDRM.

The other critical remedy to minimizing risks and costs is bringing as much of the e-discovery process as possible "in-house". Of course, this means that in-house staff must have a thorough working knowledge of the relevant processes, organizational archiving and data structure and enough technical know-how to choose and implement the right tools to support the required processes, which include (data) identification, preservation, collection, pro-

cessing, review, analysis, production and presentation.

But these components, and the process as a whole, may have variable definitions in different environments. It's important that anyone engaged in e-discovery understands the preferable way in which these components should be interpreted and executed in order to optimize the efficiency of their processes. In sections C, D, E, F, and G in Chapter 4, some very nuts-and-bolts information is presented about the reasoning and tools that need to be in place in order to ensure that the components of any e-discovery process are interpreted in the most appropriate and effective manner.

In terms of actual technical tools that can offer real impact to an e-discovery process, section H of Chapter 4 presents an array of important information about the need for a proper, legal-based search tool to drive the discovery process. Too often, e-discovery processes are compromised simply because the investigative professionals have defaulted to an inappropriate, Web-based search tool.

But when it comes to tools, it's not just a question of having the right search engine in place. High-quality search and a detailed working framework are a solid foundation on which to conduct e-discovery activities. However, the complexity e-discovery requires a nimble and comprehensive set of tools to support all aspects of the process. Appendix A provides an overview of one of the most highly regarded solutions on the market, ZyLAB's ZyIMAGE eDiscovery Platform, which bundles a powerful, specialized

search tool with a suite of technology and a fully documented best-practice methodology and working instructions. Utilizing this type of comprehensive solution is the only way those engaged in e-discovery can now ensure that their cases are aligned with expected processes and are thorough and accurate enough to stand up in court.

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

4

5

Sponsored by: **ZyLAB**®
eDiscovery & Information Management

Introduction

The discovery process involves the gathering and management of information that supports some or all phases of a business process, such as fraud investigations, financial auditing, research and development activities and more. The concept of discovery, however, is probably most associated with the legal industry. In this context, discovery is the pre-trial phase of a lawsuit in which each involved party can, through the law of civil procedure, request documents and other evidence from third parties.

In American law, discovery is wide-ranging and can involve any material that is relevant to a case, except information that is privileged or the work product of the opposing legal team. In practice, most civil cases in the United States are settled after discovery. At this point, both sides usually are in agreement about the strength of each side's case, and this realization can produce a settlement that typically minimizes the expense and risks of a trial.

Disclosure refers to the giving out of information, either voluntarily or in accordance with legal regulations or workplace rules. Some disclosures could be contrary to law, custom, or even ethics, such as the disclosure of a secret.

In the past, paper was the main information carrier, but nowadays, databases, networks, computer systems, servers, archives, backup or disaster recovery systems, laptops, personal digital assistants, mobile phones and pagers can all be considered suitable material for (e-) discovery activities. Traditional disclosure was all about

sending interested parties as much paper as possible. With e-disclosure, all information can now be disclosed in an electronic format by copying data and distributing this information or even by sharing the same information sources using secure web technology.

With the advent of so many new regulatory mandates, it's hard to find an organization that will not be (potentially) impacted by e-discovery activities. Discovery guidelines are in place around the globe, the most influential being the Federal Rules of Civil Procedure (FRCP) in the US and Part 31 of the Civil Procedure Rules (CPR) in the UK. Although other regional discovery regulations (such as those of the EU) can differ greatly in scope than the US and UK models, regional regulations are often overridden by the fact that any organization doing business, either directly or indirectly, with companies based in the US or UK must adhere to the respective US or UK disclosure standards.

Many European, Asian or Latin American companies are unaware of the far-reaching implications of the US and UK standards; that is, they are unaware until they get subpoenaed by US or UK lawyers. The ultimate solution for addressing e-discovery requirements is to implement a fully operational records management system and ensure that proper processes are in place to support it.

How many organizations can honestly say they have the proper records management infrastructure and retention policies in place to meet the requirements for any effective e-discovery solution? Perhaps surprisingly, statistics have

indicated that only around 1% of organizations are actually prepared for full-scale e-discovery activities.

As a result, if the time should come when an organization gets sued, that organization will likely jump into rapid reaction mode, immediately running to hire external bureaus or legal teams to help meet its discovery obligations. At that precise moment, not only does the (money) meter start running at an accelerated pace, but normal business processes get disturbed, management and employees get distracted and normal cash flow is compromised.

If not done correctly, discovery activities can also cause delays in court proceedings, which trigger even more organizational stress, chaos and re-allocation of resources. This situation no longer has to be the norm, though. Solutions now exist that empower organizations to prepare for e-discovery activities without having to relinquish full control to external parties.

An important component of records management is to make sure that older record collections can be recognized and organized, and the records that are no longer needed, or required by law to be retained, must be destroyed. These older documents can cause the most problems during litigation, and if they are not accounted for, any other discovery activities that occur, whether by your organization or a third party, can be rendered ineffective.

Unfortunately, few organizations are prepared to conduct these activities and the cost of organizing and managing older files can be cost-prohibitive. Regardless, many orga-

nizations realize that they must start implementing some form of records management, knowing that over the long run they cannot afford to ignore their need to handle all of their unstructured e-mail and hard disk collections. This issue is especially pronounced when one takes into account that the unstructured legacy collections of electronic information will inevitably be integrated into fast-growing collections of new types of media (blogs, IMs, external Web pages, news groups, voice mail, and so on).

Given the reality of e-discovery, the best way for any organization to prepare and optimize its capabilities, whether it plans on eventually using a third-party or not, is to integrate a comprehensive and compliant system with which to organize, control and store all of its data. A recent independent survey of corporate legal departments showed that two thirds of the respondents wanted a (e-discovery) system that was "soup to nuts," i.e. that took them from data preservation and collection all the way through production.

However, although the marketplace is certainly receptive to the cost-saving benefits of implementing such a system, several inhibitors still exist that keep organizations from fully embracing the implementation of a suitable, internal e-discovery system framework: a lack of knowledge about a system's full price/value relationship; unfamiliarity with the full scope of relevant regulations; concerns about extended liability; and fear of procedural missteps. Fortunately, a variety of affordable, straightforward, start-to-end solutions are now available that take the mystery out of

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

6

7

Sponsored by: ZyLAB®
eDiscovery & Information Management

e-discovery processes. The right system, combined with proper professional services, can enable the brunt of e-discovery work to now be performed in-house.

All told, discovery for many companies means high cost without assurances of high levels of accuracy and efficiency. The root of the problems discussed here—and the main reason pre-trial costs go through the roof—is that most organizations have no overriding concept or supporting structure in place to define and manage the relevant information that could be vital to their defense during litigation proceedings.

When information gathering and preliminary analysis start from ground zero, organizations are more prone to try to settle the case out of court, before the discovery phase. Even though it can feel like extortion, organizations often prefer to just swallow their pride and pay the high cost of a settlement because it is still perceived to end up costing less than the estimated costs and hassle of discovery and legal reviews.

In this white paper, developed in close cooperation between ZyLAB and EDRM, the rewards but also the risks of bringing e-discovery in-house are addressed and a strategy is presented to realize in-house e-discovery.

Understanding the Issues Associated with eDiscovery

CONTROL COSTS

E-discovery is a costly process, no matter how you approach it. However, with care and planning you can manage and even reduce those costs, which can be divided into two broad categories: the costs for technological systems and the costs for the people who put in place or use those systems. Careful investments in the former can, sometimes, lead to noticeable reductions in the latter.

While each stage of the e-discovery process has both people and systems costs associated with it, two stages, review and processing, generally account for the bulk of e-discovery expenditures.

Review for relevance and privilege usually is the most expensive part of the e-discovery process. Estimates vary, but the cost of reviewers appears to account for anywhere between 30 and 70 percent of companies' e-discovery, litigation or total legal budget. Whereas a number of factors influence how much a company spends on reviewers, the single most important factor seems to be the degree of control the company has over the e-discovery process. As the degree of corporate control increases, the cost of reviewers diminishes.

Review systems also can be expensive although generally they cost far less than the actual reviewers. Some review systems are internal: the company acquires hardware and software, adds those to its IT infrastructure, and maintains those systems itself. Other review systems are external with electronically stored information (ESI) loaded on an outside e-discovery service provider's system. As a general proposition, using an external, hosted review platform is more expensive. Nonetheless many factors can push an organization to choose an external system. For example, the company may lack the wherewithal to host the data itself, certain strategic considerations might require utilizing an external system, or the organization may feel constrained by established cost allocations that don't allow for the integration of an internal system.

Processing often is the second-most expensive part of the e-discovery process and consists of two main components:

- Reduce the post-preservation and post-collection body of ESI to a more manageable size. This activity is typically achieved through a series of exclusionary exercises that remove from further consideration ESI that meets pre-determined criteria. These criteria can include file types, file creation or modification dates, storage locations, or the lack of key words. This activity is usually the less expensive of the two processing components.
- Convert ESI from the formats in which the data existed when it was identified, preserved (if necessary), and collected to one or more formats that make review easier for the reviewers. In the past, this process meant:
  - Converting nearly all ESI to TIFF images

  - Extracting some (if not all) of the metadata associated with the ESI
  - Extracting some (if not all) of the text associated with the ESI
  - Loading those three sets of information into a review system.

Now, though, conversion is more of an exercise in providing reviewers either with access to files in their original ("native") format or with access to ESI that has been converted to a different format but is presented in a ("near native") way that mimics the native display. This activity is often the most expensive component of processing.

CONTROL RISKS

With each stage of the e-discovery process comes with risks as well as costs. The most obvious, often most serious, risks arise early in the process: at the identification, preservation and collection stages. These errors can be irreparable. If, for example, critical files are destroyed after a litigation hold has been implemented and before discovery copies have been made, courts are showing an increasing willingness to impose severe penalties on the offending party. These penalties have included:

- Substantial fines
- Adverse inference instructions, where a jury may be told, for example, that it must assume that the offending party intentionally destroyed the missing information to cover up its bad acts

Bringing eDiscovery In-House: RISKS AND REWARDS by George J. Socha, Jr., Esq.

8

9

Sponsored by: **ZyLAB**® eDiscovery & Information Management

- Dismissal of the offending party's claims or defenses or a finding of liability against the offending party

By contrast, errors at later stages – processing, review and production, for example – usually can be corrected. The resulting penalties, if any, tend to be far less serious.

Less obvious risks accompany the information management stage. The goal is to put and keep your electronic "house" in order: If your electronic house is in disarray, you run a high risk of being unable to identify, preserve and collect pertinent ESI. Even if you are able to find pertinent ESI in a disorderly electronic house, doing so can take much more time and cost much more money than would be the case if your ESI were better organized.

Other risks include:

- Loss of control over ESI if the ESI gets disseminated to multiple outside counsel and e-discovery service providers
- Inconsistent processing, review and production of ESI if the company does not avail itself of a centrally controlled repository

RECOGNIZE BENEFITS AMONG A MINEFIELD OF NEGATIVES

E-discovery does not just impose costs and pose risks. A well-designed and deployed e-discovery process can offer corporate personnel a wide range of benefits, which can include:

- Faster, less expensive and more accurate identification of potentially relevant ESI
- Earlier and more well-informed assessments of the circumstances surrounding disputes, which can lead to more effective strategies for handling disputes
- Tighter control over the organization's ESI while the dispute is active
- More consistent handling of ESI across disputes
- Increased opportunities to reincorporate ESI into the organization's records management systems following final resolution of disputes
- Reduced costs, especially at the review and processing stages
- Reduced risks, especially at the identification, preservation and collection stages

Addressing the Problem: Reducing eDiscovery Costs & Risks

ADOPT A RELEVANT AND WORKABLE FRAMEWORK SUCH AS EDRM

To better contain the costs and control the risks of e-discovery, it helps to work within a recognized and widely used framework. Having a common framework provides a standard language and a set starting point for determining what to do.

In today's e-discovery arena, the two common frameworks are:

- The Sedona Principles (http://www.thesedonaconference.org), which offer a framework for the analysis of legal issues associated with e-discovery.
- The Electronic Discovery Reference Model (EDRM) (http://edrm.net), which provides guidance on how to carry out the steps involved in e-discovery and offers ongoing information about the developing technical standards for e-discovery

TAKE EDISCOVERY IN-HOUSE

For a certain cadre of corporations, one of the best ways to address the e-discovery problem is to begin bringing e-discovery in-house. This concept relies on getting, or taking back, control of the e-discovery process. With direct control can come reduced costs, reduced risks, less internal disruption, more consistent actions and results, and greater predictability.

How an organization takes or resumes control varies greatly from one organization to the next. Taking or resuming control also means different things at different stages of the e-discovery process, as outlined in the subsections below.

Information Management

In this context, information management means getting your electronic house in order, which ought to be a proac-tive step. Information management also can involve the initial generation of ESI and should involve its final disposition.

Effective information management requires close collaboration among a wide range of internal personnel (legal, IT, records management, information security, human resources, audit, business units, and so on) and involves modifications to and enhancements of internal systems.

If these tasks are accomplished, an organization ought to be in a much better starting position when it faces an e-discovery challenge. The organization will be better able to understand what ESI it has, where that ESI is located, who has control over it, and what will be done with it throughout the e-discovery process.

Identification

Identification is the first reactive e-discovery step and entails locating potential sources of ESI as well as determining the scope, breadth and depth of that ESI. Reasons to move identification in-house include:

- Greater control over the processes used to identify ESI of potential interest
- Greater familiarity with the organization's data, data sources, people who generate the data, people who work with the data, and so on
- Ability to conduct a narrower, more focused identification process
- Greater chances of following a consistent process over

multiple matters

- Less disruption to custodians – the people whose ESI is of potential interest
- Less disruption to the people who maintain the IT infrastructure
- Lower identification costs.

Reasons to leave identification in the hands of others include:
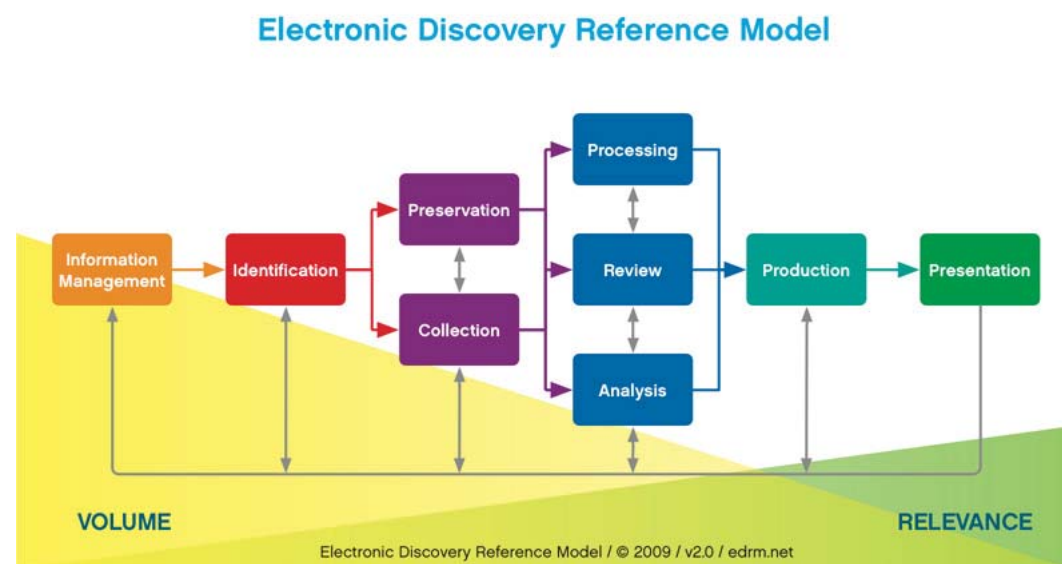
- The organization lacks the resources to engage in the identification activities itself
- It does not have personnel with appropriate training and experience to reliably identify ESI of potential interest
- It cannot expect to complete the work on time with current staff
- The issues in the matter are such that use of the organization's own personnel to identify ESI would be viewed as, at minimum, a lapse of judgment.

Preservation

Preservation consists of ensuring that ESI is protected against inappropriate alteration or destruction.

Reasons to move preservation in-house, or leave it in the hands of others, match those listed under identification. In addition, moving preservation in-house can reduce the likelihood of over-preservation.

Electronic Discovery Reference Model

Electronic Discovery Reference Model / © 2009 / v2.0 / edrm.net

Collection

Collection means gathering ESI for further use in the e-discovery process (processing, review, and so on). At times, preservation and collection are one and the same; at other times, they are two distinctly separate activities.

The reasons for bringing collection in-house or leaving it to outsiders match those for preservation.

Processing

Processing has two major purposes:

- To reduce the volume of ESI that has been preserved and collected and has been teed up for review
- To convert the ESI, if necessary, to forms more suitable for review and analysis

For the most part, corporations have left processing activities to others.

Review

Review is the evaluation of ESI for relevance and privilege. This step almost always takes place after identification, preservation, collection and processing, and before production.

Review has two components: the platform used for review of ESI and the reviewers themselves.

Most corporations leave review in the hands of their outside counsel. As such, the corporations will allow the outside counsel to dictate the platform to use and determine who conducts the review.

Some corporations have begun to make other arrangements and are now telling outside counsel what review platform to use, who to use as reviewers, or both.

The primary considerations when deciding who controls review tend to be all of the following:

- The ability to reuse both the processed ESI and the review calls
- Greater consistency in review calls
- Greater control over the distribution of the company's data
- Greater control over costs.

Analysis

Analysis is a more in-depth evaluation of ESI for content and context, including key patterns, topics, people and discussions. Efforts spent on analysis, as opposed to review, can lead to a better understanding of the facts of the case, which in turn can help direct the handling of a dispute.

Both in-house and outside personnel ought to be more focused on analysis than generally is the case today. As discussed in the "'faith-based' e-discovery" section later in this document (C1), in-house personnel should be using analytical techniques starting as soon as they determine the need to engage in some form of e-discovery activity. These techniques enable users to better understand what

"Faith-based e-discovery means e-discovery activities undertaken by an organization with the implicit and untested belief that the activities will deliver the desired result."

"The other side reviews the materials it has received and on the basis of that review concludes that a substantial number of files that they think ought to have been produced were not."

they need to do and what their data tells them about the dispute at hand. Outside counsel should be using analytical techniques at every stage of the e-discovery process to help keep themselves on track, to avoid unpleasant surprises, and to better understand the substance of the data they are handling.

Production

Production means delivering ESI to others in appropriate forms, using appropriate delivery mechanisms.

Often, but not always, this process includes converting materials to a TIFF format (if that has not already been done); affixing an identifying number to each TIFF image; and redacting privileged or confidential information from the TIFF images. Production can also mean associating identifying numbers with files kept in a native – or near-native – form for production, as well as tracking what materials are being produced, by whom, to whom, for whom, at what time, and for what reasons.

In general, production is an activity that is undertaken by either outside counsel or by services providers working on their behalf.

Occasionally, corporations take responsibility for this activity, particularly with small productions or where the company has elected to be its own e-discovery general contractor.

Presentation

Presentation consists of displaying ESI before audiences (at depositions, hearings, and trials), especially in native and near-native forms, to persuade or to elicit further information. This activity is almost always the domain of outside counsel.

GO BEYOND A FAITH-BASED APPROACH

While bringing e-discovery in-house can help an organization better control its e-discovery costs and risks, simply the fact of bringing the processes and tools in-house is not sufficient. Organizations also need to move beyond "faith-based" e-discovery and on to approaches that start from sound empirical bases and involve regular testing and evaluation.

What Does Faith-based Mean?

Faith-based e-discovery means e-discovery activities undertaken by an organization with the implicit and untested belief that the activities will deliver the desired result.

Organizations have pursued faith-based e-discovery at every stage of the e-discovery process. Below are four real-world examples, all of which share a common failing: Everyone makes assumptions about how a part of the overall e-discovery process will operate and what it will yield, but no one tests the assumptions to determine whether they are valid:

Example 1: Information Management, Identification, Preservation and Collection:

1. Company A decides that to better manage its messaging system while simultaneously reducing its e-discovery costs, it would like to implement an e-mail archiving system.

2. The provider's sales personnel assure Company A that the e-mail archiving system will capture all the messaging information that the company is going to need for discovery.

3. Accepting this assertion as reliable, Company A purchases and installs the system and begins using it.

4. Company A is sued. Company A issues a legal hold but, assuming that the e-mail archiving system preserved all the messaging information of potential relevance to the lawsuit, does not include e-mail within the scope of materials covered by the hold.

5. Company A collects messages from the archiving system and sends them to a processing provider to be loaded onto a review platform. Outside counsel reviews the messages and produces a subset to opposing counsel.

6. After examining the produced messages, opposing counsel demands that Company A produces the path information for each message as well as calendar entries for a group of key Company A employees.

7. Upon examination, Company A discovers that the archiving system did not retain path information, did not archive calendar entries, and altered e-mail creation dates and times.

8. Because months have elapsed, there is no way for Company A to go back and retrieve the missing information.

9. Motion practice ensues, and the story just keeps getting worse.

Example 2: Review

1. Company B has recently been sued over a set of actions it took that it never envisioned would lead to litigation. Wanting to avoid preservation disputes, Company B preserved a much broader range of data than it ever expects to produce to plaintiffs. That data has been collected, processed and loaded into a review platform.

2. Outside counsel needs to review the data for relevance and privilege. Lawyers for the outside firm decide to prepare a list of key words, and they meet in a conference room, armed with the complaint, their answer, and the limited information they have gathered through interviewing company personnel. Referring only to these materials and talking among themselves, the outside lawyers draw up a list of 150 words or phrases, believing that a search of the data using these key words and phrases will give them the right set of documents for their staff to review for relevance and privilege.

3. The outside firm instructs the provider to make one pass through the data using those words and phrases. Any file containing at least one word or phrase should be teed up for review. Any file containing none of the words or phrases should be set aside, presumed to be irrelevant.

4. The provider complies. The firm reviews the first set of files and produces a subset of those files to the other side.

5. The other side reviews the materials it has received and on the basis of that review concludes that a substantial number of files that they think ought to have been produced were not.

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

14

15

Sponsored by: **ZyLAB**®
eDiscovery & Information Management

6. Company B's outside counsel insists their process was appropriate.

7. Motion practice ensues. Plaintiffs ultimately obtain an order requiring Company B's outside counsel to re-do the entire review.

Example 3: Processing and Review:

1. Company C needs to provide a regulatory agency with data within a short timeframe. Having never before dealt with a major e-discovery challenge, Company C relies on the advice of its outside counsel.

2. Outside counsel gets a proposal from just one provider, with a fee estimate of between $135,000 and $850,000. A more precise estimate, Company C is told, simply is not possible.

3. Outside counsel begins feeding data to the provider, which processes the data as it receives it and loads it to an on-line review platform.

4. As the volume of data grows beyond anything initially anticipated, outside counsel tells the provider to keep processing and loading the data.

5. The provider sends outside counsel written status reports twice a week. The reports show the increases in volume of data handled but do not show any total of fees incurred. Outside counsel does not send the reports on to the client. The provider also sends invoices, but none of those shows the total fees either.

6. At the end of the third week, the provider finishes processing. Shortly after that the law firm finishes reviewing the data

and produces a portion of the reviewed data to the agency.

7. Only then does the law firm send the accumulated invoices to the client. The total is $2.7 million, far beyond the anticipated high end of $850,000.

Example 4: Processing and Review Redux

1. Continuing from Example 3, the materials sent by the law firm to the provider included, among other items, nearly 1,000 Microsoft Access databases.

2. The provider asked the law firm how it would like the database files to be handled. The law firm replied that it did not know and sought guidance from the provider.

3. The provider recommended TIFF'ing the database tables. The law firm accepted the provider's recommendation. No one discussed how those images might be used.

4. The law firm received many hundreds of thousands of TIFF images of database tables, information which it was not able to use in any productive fashion.

5. The provider charged approximately $1 million for TIFF'ing the databases. Ultimately, neither the law firm nor the company paid for that work.

Moving Beyond Faith Based eDiscovery

Moving beyond faith-based e-discovery is easy – conceptually. In practice, however, the move can be difficult to make. Three key steps are involved with enabling an organization make the transition: use the scientific method, use iterative approaches, and measure and evaluate what you do.

Use the Scientific Method

Formulate a hypothesis and test a hypothesis, reformulating and retesting as needed.

With Example 4, your hypothesis might be that the best way to process, review and produce an Access database is to print the database table as TIFF images.

After formulating a hypothesis, test it. Start with a single database, and select a single table from that database (assuming it has more than one table). Generate some TIFF images (five images, for example).

Then, evaluate what you have done and what it has given you. Ask whether a reviewer looking at the test TIFF images will be able to make a decision about the relevance of the database. Ask, as well, whether a recipient of the TIFF images would consider them to be a reasonably useable form of the database. Ask whether, if you were on the receiving end, you would accept TIFF images in lien of the database.

Use Iterative Approaches, with Iterations Informed by Knowledge Gained from Previous Efforts

Use an iterative approach, rather than a single pass through the data. Most steps in any e-discovery process need to be started long before anyone involved has enough information to develop a complete picture about what ought to be done. Many times, as well, companies find themselves undertaking e-discovery activities that they have not previously had to address.

As a result, one pass through a process often is not sufficient. It may be that only after multiple passes do you reach a point where you feel comfortable that the product of your work is reliable.

In Example 2, outside counsel relied on a single pass through the entire body of the data. Had they used an iterative approach, counsel would have had a much better chance of honing in on the most important data, conducting a more focused and nuanced review, and enhancing the chances of delivering a defensible production.

An iterative approach to the problem in Example 2 could look like this:

1. Based on the limited information available to outside counsel at the time, outside counsel identified between one and no more than, say, five Company B employees who appeared to be at the core of the dispute.

2. For those key employees, outside counsel collected as much information as was available from three locations: those people's e-mail accounts, the hard drives of their desktop or notebook computers, and the locations on the company network designated as places where people could store their files (often called "file shares").

3. Outside counsel deployed a small number of skilled people to go through that data, using a range of analytical tools, to identify concepts key to the dispute.

4. The same people then converted the concepts to key words or phrases and tested the efficacy of those words and phrases against that first body of data, data from another

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

16

17

Sponsored by: **ZyLAB**®
eDiscovery & Information Management

small group of custodians, and data from some outside sources.

5. Based on the results of the testing, they refined and retested the words and phrases, until they reached a point where they were comfortable that those words and phrases, when used with a larger body of data, would be reasonably effective.

Generally, an iterative process akin to the one described above can be accomplished quickly and at a much lower cost than that of reviewing large amounts of data that never really were likely to be of any consequence to the dispute.

Measure What You Do and Evaluate What You Measure

All too often, no one involved in an e-discovery exercise has any real idea how much time the exercise will take, how much it will cost, or how much data it will involve.

As a result, it becomes difficult, if not impossible, to establish reliably timelines, set achievable targets, develop realistic budgets, and manage expectations. If you have a sound understanding of how long an e-discovery process will take, how much it will cost, and how much data will be involved, you should be able to do a reasonably good job of meeting all these stated objectives.

Reaching this point takes time, however:

1. You need to start measuring time, money and volumes – something that almost no one seems to do today. Some-

time, you may be able to take the measurements yourself. At other times, however, you will need to have others take the measurements for you. But developing these markers will only work if everyone takes measurements using the same yardsticks.

2. You then need to take enough measurements to have a meaningfully large set of data points. One data point tells you nothing because you have no idea where the second one might land. Two data points are little better. Only with three data points can you begin to see any patterns; with three data points, for example, you know whether you have a line or a curve.

3. You need to evaluate the data. You can begin to determine how long a process tends to take, which approaches appear to be more effective, which are more efficient, and so on.

Use an "Investigative" Approach, Not Just an "Exclusionary" One

Using an investigative approach to e-discovery, as well as an exclusionary one, can help you arrive at the most important data in a more expedited and cost-effective fashion than if you rely solely on an exclusionary approach.

Defining an Exclusionary Approach

The exclusionary approach to e-discovery is a common practice today. The essential idea is to set aside classes of ESI that do not appear to be of consequence. The remaining materials are deemed to form the body of potentially pertinent ESI that will be reviewed for relevance and

privilege, hopefully within the available budget and time constraints.

A number of elimination criteria are used to exclude presumptively inconsequential ESI. Each of these criteria can be applied in a positive fashion, where only the ESI meeting the criterion will be kept in the workflow for further consideration, or in a negative fashion, where the ESI matching the criteria are set aside.

At times one party, acting unilaterally, may define and deploy the criteria. At other times, the opposing parties work together to develop the criteria.

Types of commonly used elimination criteria include:

- **File types**: Sometime file extensions are used, such as ".doc", ".xls" and ".pst". Sometimes file headers are read to determine file types. Often, certain files types, such as ".exe" files, are assumed not to be of interest. Other types, often database, audio or video files, are set aside as exceptions that receive no further handling.
- **Date ranges**: Files created or last modified before or after certain dates might be set aside.
- **Custodians**: A data custodian might be added to a list of people whose data should be evaluated or to a list of people whose data need not be examined.
- **Data locations**: A data storage device might be identified as one to set aside, or a geographical location, perhaps a satellite office, might receive the same treatment.
- **File size**: Files that fall outside certain size limits may be set aside, either because they are so small that they are

deemed unlikely to contain any information of interest, or because they are considered too large to handle properly.

Defining an Investigative Approach

The investigative approach to e-discovery is less commonly used today, although in earlier days it was a frequently used practice, grounded as it is in traditional computer forensics. The goal of the investigative approach is to focus as early as possible on the ESI that appears to be most important. Often the investigator places greater weight on the content of the ESI than on its form, seeking, for example, communications that discusses a specific transaction or database records that show a certain behavioral pattern.

The initial goal is to find any piece of information the meets the sought-after parameters. If even a single piece of information is found, then the objective becomes one of using that piece of information as a foundation on which to expand the search. Expansion might mean instigating efforts to find more of the same or attempts to follow whatever path the initial piece of information points out.

If no information is located, the investigator adjusts the parameters, either widening or refocusing them, and tries again. If the initial search returns too much information, the investigator narrows the focus.

> "Even worse can be a workflow that doesn't enable you to retrace your steps, which means if a major omission is not identified until late in the process there is no way to correct the problem."

> "Only once did I encounter a situation where no organizational principles were in place: The president of a smaller company kept all his working files in about half a dozen boxes."

Understanding the Emphasis on Investigative over Exclusionary Approaches

Acknowledge the pros of an exclusionary approach. For several years now, the exclusionary approach has been dominant in e-discovery circles. As a result, many people who got involved in e-discovery in the last five or so years are most familiar with the exclusionary approach, find it to be most comfortable, and have the easiest time executing it. Unless these people are pushed toward an investigative approach they automatically pursue an exclusionary one.

The exclusionary approach also is easier to define, which makes it more attractive to litigators. Parties often feel that if they can come to an agreement about, for example, the file types to be reviewed, then they will have reduced potential areas of conflict.

Recognize the cons of an exclusionary approach. The exclusionary approach often operates more as a blunt instrument than a precision tool. By selecting or de-selecting materials based on secondary characteristics – file type, date range, and so on – instead of based on the criteria you really care about – such as content that reflects an important decision – it is easy to unintentionally set aside specific files that matter considerably. For example, if you chose to examine only files with e-mail or office file types (.msg, .doc, .docx), you would be overlooking relevant content stored in a word processing file whose name had been changed from, for instance, "important_file.doc" to "important_file.doc.old".

As datasets grow in size, the chances of inadvertently excluding relevant materials grow even more. In part, this risk is simply a function of volume; however, the fact also remains that the larger the dataset, the greater the temptation to do something, anything, to reduce it to a more manageable size.

Another danger with pursuing a predominantly exclusionary approach is that the exclusionary approach often is used in connection with an e-discovery work flow that delays in-depth examination of content until late in the process. By the time someone realizes that a potentially important body of data has been overlooked, there may no longer be enough time or money to return to that data. Even worse can be a workflow that doesn't enable you to retrace your steps, which means if a major omission is not identified until late in the process there is no way to correct the problem.

Understand the pros of an investigative approach. If well planned and executed, an investigative approach offers the possibility of finding and focusing on highly relevant information early in the process. When this early identification happens, the investigative approach can become a means of bringing a matter to resolution earlier, at a lower cost, and with less disruption to the organization.

Even if an early resolution is not the outcome, the investigative approach should start feeding relevant data to decision makers sooner than is the case with an exclusionary

approach. This capability gives them a better understanding of the problem, a better idea of directions in which to take the handling of the matter, and a better ability to direct further investigative efforts along productive paths.

Don't ignore the cons of an investigative approach. The investigative approach has two primary drawbacks, depending on your perspective:

- It can be harder to carry out, at least initially, because people are less familiar with it.
- It can mean reduced revenues for those who make money from reviewing ESI for relevance and privilege.

March Full Speed into the Past

By going beyond a faith-based approach and pursuing an investigative as well as an exclusionary approach, one marches full speed into the past.

Remembering "Once Upon a Time"

Twenty years ago, when e-discovery was unheard of to most litigators, dealing in discovery meant dealing in paper (and interviews, depositions, and the like, of course). Dealing in paper, in turn, sometimes meant going to warehouses full of retained materials.

A young attorney or a green paralegal, having never been through a warehouse of paper before, might have been tempted to start at one corner of the warehouse and work through the materials box by box, folder by folder, page by

page. That approach, if attempted, faced almost certain failure; you ran out of time or money long before you got all the way across the warehouse.

Usually, however, someone more seasoned made sure this did not happen. The first step was to gauge the lay of the land. How was the warehouse set up? Was there shelving? Was it marked? Were there boxes on the shelves? Were they marked? On the outside or the inside? How were materials organized inside the boxes?

Sometimes we went in, clipboard in hand, and drew the warehouse layout on graph paper. Once in a while, we photographed the space. Wherever possible, we found someone who worked in the warehouse, someone who could describe to us the organizational principles involved.

Only once did I encounter a situation where no organizational principles were in place: The president of a smaller company kept all his working files in about half a dozen boxes (he did not have any retained records). In one of our first meetings with him he had the boxes brought into a room for us. He then upended the boxes, dumping their contents into a single pile, and got down on his hands and knees and shuffled through the paper. That was, his colleagues insisted, the way he always searched his files.

By taking the lay of the land, we gained the ability to get an overview of what we were up against. We could plan what to tackle when, which enable us to develop a preliminary cost estimate. We were able to put materials into batches and prioritize the batches – "these 60 boxes

"Using text mining, text analytics, auto-categorization and similar tools and techniques, you can prepare specialized sets of review materials."

"Legal professionals...often default to in-house variants of common Web search tools. Those tools, however, are not optimized for the types of activities associated with e-discovery."

are insurance records unrelated to the dispute; these 20 boxes are marketing materials we will have to go through in detail; these 10 boxes are accounting records that may be important but that no one in the room at the time has the expertise to evaluate…" – and so on.

Embracing Today

Today, the old ways seem to have been forgotten. Or, at least, most people never translated them into the world of electronic materials. Instead of trying to learn the lay of the land, counsel pursue an approach akin to zigzagging through the warehouse:

1. Preserve data, often by asking custodians to set aside or forward the information they think is important.

2. Collect the set-aside or forwarded data.

3. Process the collected data.

4. Formulate a set of key words, often without looking at any data first.

5. Use those key words for a one-pass search through the data.

6. Send on the results for review for relevance and for privilege.

7. Produce the file designed by the reviewers as relevant and not privileged.

8. Never go backward: Do not evaluate the processed files deemed to be irrelevant in order to determine whether they actually contained information that mattered. Do not evaluate the files that did not contain key words, and do not feel compelled to evaluate the materials not identified by custodians as relevant.

9. Do not take any steps to refine or refocus your efforts; having not gone backward, it isn't necessary.

Formulating a Better Approach

The time to return to the old "warehouse" approach is here. Start with an overview of the ESI. Prioritize materials, and assess your progress as you go. Here are some of the key, and often ambitious, steps:

1. Try to get your electronic house in order (akin to putting shelving in the warehouse, putting paper files into boxes, redwells and folders, and labeling everything):

• Set up systems to archive e-mail and other files

• Create maps of systems and then keep the maps current

• Create standard operating procedures and follow them

• Make it easy for users to store and retrieve information using systems that also allow the organization to readily determine what information is stored in what location and to get at that information

2. Implement a more formal, structured, systems and policy-based approach to litigation holds, preservation and collection

3. As soon as you begin identifying custodians or other sources of electronic information, start looking at and evaluating their ESI

4. Based on that evaluation, try to determine whether ad-ditional identification efforts are needed, where else to go to preserve ESI, and whether additional information needs to be collected from previously preserved ESI

5. Also based on that evaluation, begin making judgments about how to handle the dispute or other issue that has caused you to conduct e-discovery

6. Preserve more data, if necessary

7. Evaluate additional data, if necessary

8. Repeat as needed.

Move to Higher Octane, Analytics-Fueled Review

Drawing on analysis to enhance review can take yet other forms. Start by using more advanced searching, analytical and organizational technology to create views of the ESI that you cannot get with more basic Boolean searches and row-and-column reporting of results. Using text mining, text analytics, auto-categorization and similar tools and techniques, you can prepare specialized sets of review materials.

Once materials are better organized by content, send those highly specialized materials to specialists for review. The specialists might be lawyers experienced in a particular area of the law. They might be engineers, doctors, accountants, statisticians – anyone with an in-depth understanding of the contents of the particularized set of ESI.

At the same time, consider putting specialized file types together. If you have databases that require review, hand them to someone adept at working with that type of data-base. If you have complex spreadsheets where computational skills are needed to understand them, send those to someone skilled at working with those types of materials.

Use in-house people to review, or pre-review, materials. Often they will have a better understanding of the content than anyone outside can be expected to develop and will be able to put that expertise to good use, improving the quality of the review while reducing the costs.

Deal with Complex and Less Common File Types, Don't Just Treat them as Exceptions

Complex files need to be addressed, not avoided. Generally this means they need a level of processing not required for more straight-forward files.

Complex files include, but are by no means limited to, the following types:

• PST

• ZIP

• SharePoint

• Oracle

• Livelink

• SQL

Conversion of the files may be required. Sometimes connectors are available to allow more direct and highly-con-

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

22

23

Sponsored by: ZyLAB®
eDiscovery & Information Management

"People involved in criminal activities (such as fraud) don't want to be in the top 10 of a search engine result list, so they use advanced techniques to hide."

"When done right, all or parts of e-discovery can definitely be brought in-house; by doing so, you can save tremendous amounts of money."

trollable access to the content. At times, the use of APIs or other toolkits is the best approach to take.

Often files are not initially searchable. Some, such as ZIP files, can be processed in ways that make the content readily searchable. Others, such as bitmaps, audio and video files, call for the deployment of entirely different search technologies (if those technologies are even available).

Multi-lingual files present their own challenges. Most obvious among the challenges are:

- How to process files that are in a language other than, for example, English
- How to process files that contain content in more than one language
- How to search the contents of those files, once they have been processed
- Who conducts the search
- Who conducts the review.

Know the Differences Between Legal Search and Web Search Technologies

When legal professionals require advanced searching capabilities for e-discovery and legal activities, they often default to in-house variants of common Web search tools. Those tools, however, are not optimized for the types of activities associated with e-discovery. This inadequacy is due in large part to fundamental differences between the capabilities of Web search engines and the search func-

tionality and approaches needed to support the strategic requirements of legal, law enforcement and intelligence applications.

Understanding the Depth of Indexing

With a Web search engine, you may not know exactly what data is in your index, and more specifically, what data is not in your index.

Most search engines use a "tokenizer" to enhance the searchability of data by removing punctuation and noise words, identifying words, and determining character set mappings (for foreign languages). This type of capability enhances your ability to perform the necessary full-text indexing of all relevant data. Of course, Web appliances can index for you, but their reporting and auditing functions may not match the standards that you, your opponents or regulators expect or require.

A Web search appliance may only keep the 20,000 most relevant files in its index for a particular occurrence, making the search engine of limited value for e-discovery.

Many Web search technologies cannot index documents that consist of compound documents (e.g. ZIP and PST), bitmap data, multimedia documents, older electronic file formats, and encrypted files. If a legal search program runs into these types of documents, it should either separate them through a culling process or automatically include additional processing to make such files fully searchable.

To be adequate, a search technology should be able to index the full text of files as well as extract document and file properties and make them searchable.

Understanding the Depth of Search Results

Those searching in legal or law enforcement environments need to find all potentially relevant files – or at least as many as they can find before they run out of money, run out of time or find the files that really matter.

Typical Web search engines are optimized to find only the most relevant files, not all relevant ones. With Web search engines, most companies and organizations place a premium on being found as close to the top of search list as possible. Experienced users have become savvy at manipulating search engine optimization techniques to enhance high rankings. This level of sophistication works in both directions. People involved in criminal activities (such as fraud) don't want to be in the top 10 of a search engine result list, so they use advanced techniques to hide their documented activities and avoid appearing in any search list.

Understanding the Range of Functionality

Moreover, these investigators require different tool functionalities to quickly and efficiently navigate and review relevant sets of files.

Web search engines use many optimizations to continually perform real-time indexing of the Web. These optimiza-

tions, however, come at a price:

- Files in non-standard formats may not be found
- Long files may require a lot of time to review
- The processing of complex queries may be very slow (or even impossible)
- Hit highlighting and hit navigation are often not available or operate too slowly
- After files are found, tagging them is not possible
- Files cannot be exported in a format required by opposing parties, regulators or other recipients of the data.

Conclusion

When done right, all or parts of e-discovery can definitely be brought in-house; by doing so, you can save tremendous amounts of money. However, when done incorrectly, the damage can be tremendous.

Many horror stories exist in which a deadline is missed, vital information got destroyed and (parts of) the process was not done right. In some cases, these errors can be corrected, such as in the review or production phase, but if information is missing due to a badly implemented legal hold, identification or collection of information, then you can no longer correct these errors, and fines are often the only result.

Nothing is more embarrassing and more costly than to be sent back by the court to redo discovery work. Often this work has to be completed in an extremely short time

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

24

25

Sponsored by: **ZyLAB**®
eDiscovery & Information Management

interval, and in many cases, you may have no other choice then to involve expensive third parties to do the work. Here, all efforts to bring e-discovery in-house have failed, and not only has the investment in tools, resources and labor been a waste, you are also confronted with extra costs as a result of the external processing.

Bringing e-discovery in-house is not the "end all/be all" solution for the minimizing of e-discovery costs and risks; rather, the final solution lays in the implementation of records management systems and proper training of

your employees in the "arts" of records management and information destruction. For legally sensitive archives such as e-mail, HRM files, project files, clinical evidence, maintenance records and many more such repositories data must be destroyed when you are allowed to destroy it. Archive and organize your data, especially confidential and privileged information. Do not keep completely unstructured collections that hold legal risks. Unstructured collections of data that need to be processed directly into an e-discovery pipeline will always be around, but there are

The Confluence of E-mail, Records Management, E-Discovery and Knowledge Management

many other collections on which the application or records management principles will tremendously decrease the pain and the cost of discovery.

When you do select a records management or e-mail archiving tool, make sure that you select one that can implement not only a legal hold, also the complex legal searches that are required in an e-discovery process. If you cannot implement the "legal" search requirements as discussed in this white paper, than you should be prepared for very high conversion cost before you can access your data in records management repositories.
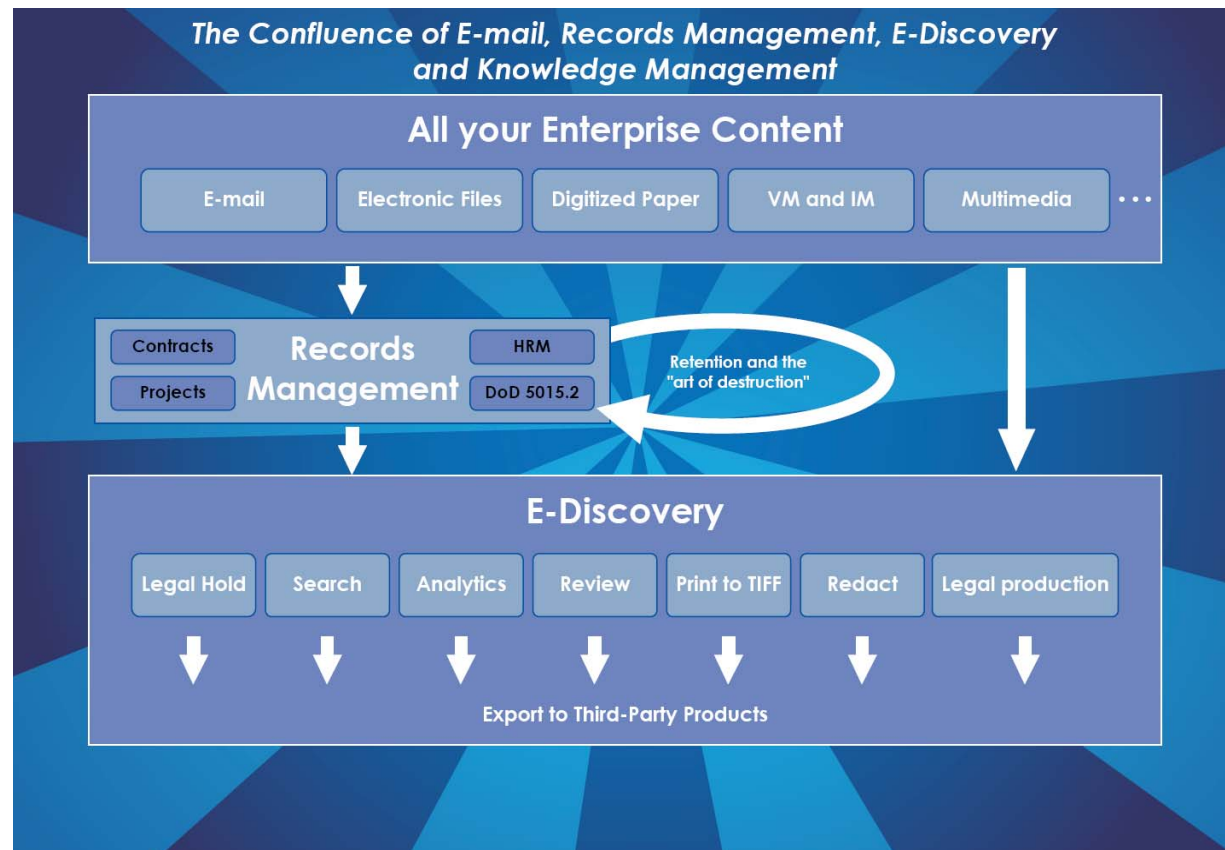
As an additional benefit, after all your data is properly organized and structured and non-relevant legacy data has been removed, your data is more searchable and the real knowledge management can start.

Therefore, you must not only buy "point solutions" or "IT-Tools", you should look for a platform that includes records management, e-mail archiving and discovery tools, as well as a fully documented methodology on how to implement different components of the e-discovery process in-house. Forms, quality control, chain of custody records, audit records and inclusion and exclusion reports are all key features worth considering.

When you do select a partner, select one that will teach you how to deploy these functionalities. An intuitive and solid platform is also required, and it needs to be properly installed, deployed and configured. Your employees need solid training and hands-on assistance. With these sup-

port functions in place, you are positioned to start implementing 100% of the work in-house. The savings can start now.

But remember: although bringing e-discovery in-house saves money, it but also comes with costs. These costs go beyond the expenditures for the actual tools; they also include real costs in terms of your time to understand, learn and implement solid in-house e-discovery procedures. ■ Authored by George J. Socha, Jr., Esq.
© ZyLAB 2009-2011. All rights reserved.

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

26

27

Sponsored by:

ZyLAB®
eDiscovery & Information Management

"In the 2011 Magic Quadrant for E-Discovery Software, Gartner describes ZyLAB as highly referenceable, extremely stable, and with a loyal client following."

Gartner Recognizes ZyLAB

Gartner, Inc. (NYSE: IT) is the world's leading information technology research and advisory company. The company delivers technology-related insight necessary for its clients to make the right decisions, every day. It is an honor for ZyLAB to be included in Gartner's series of Magic Quadrant and MarketScope reports.

ZyLAB has received numerous industry accolades and is one of the few companies to be positioned as a Leader in Gartner's "Magic Quadrant for Information Access Technology" for the last three consecutive years. In addition, Gartner has given ZyLAB the highest rating ("Strong Positive") in its "MarketScope for E-Discovery and Litigation Support Vendors" for the past three years, a "Promising" rating in its "MarketScope for Records Management", as well as a "Visionary" rating in its 2011 "Magic Quadrant for E-Discovery Software".

In the 2011 Magic Quadrant for E-Discovery Software, Gartner describes ZyLAB as highly referenceable, extremely stable, and with a loyal client following. The analysts state, "ZyLAB is a veteran of the information retrieval business that now specializes in e-discovery. It has a very broad range of language capabilities and a wider geographic distribution than the other vendors. ZyLAB functionality is equal to any of the market leaders, and it should be considered alongside them."

ZyLAB is among an elite sub-group of vendors that fully addresses the left-side and right-side of the Electronic Discovery Reference Model (EDRM) as well as providing information management. The Magic Quadrant report states, "Thanks to its long heritage in search and information retrieval, ZyLAB's capabilities include identification, collection, preservation, processing, review, production and ECA, based on strong textual analytics and other semantic technologies."

ZyLAB's Universal Approach to Managing Data

ZyLAB straddles the convergence of information management and eDiscovery to keep your content assets (and liabilities) in order and to cost-effectively mine them when an investigation ensues. For nearly 30 years, ZyLAB has been working alongside professionals in the litigation, auditing, security and intelligence communities to develop the best solution for investigating and managing large sets of information.

Today, ZyLAB features modular architecture and licensing that allows flexibility for clients in terms of the scale and sophistication of the solution. The robust ZyLAB Information Management Platform may be configured to the specific needs of corporations, law firms, law enforcement departments, and government agencies through limitless combinations of our companion bundles and modules. Additionally, ZyLAB is one of a select group of vendors to also deliver robust solutions for enterprise information

"ZyLAB software is at the convergence of a response to a specific eDiscovery and the implementation of comprehensive compliance and litigation readiness systems."

"ZyLAB eDiscovery software leverages automation and best practices to collect an exhaustive data set and then methodically whittle it down to an optimized review set."

management, Department of Defense- and Sarbanes-Oxley-compliant records management, and myriad archiving solutions.

The **ZyLAB eDiscovery & Production System** and five other specialized systems have been pre-configured according to the best practices for the legal and business requirements of today. The system is modular and incorporates relevant components from two of our six pre-configured, specialty systems, namely:

- **ZyLAB Compliance & Litigation Readiness System** supports the creation, administration, maintenance, archival and disposal of millions of dynamic and static files – from email, to legal agreements, to voicemails, to customer and employee records, and any other content archives. The ZyLAB Compliance & Litigation Readiness System fosters enterprise-wide corporate governance by helping you to implement practical, day-to-day records management policies, easily enforce their compliance, and quickly access information when you need it.

- **ZyLAB Enterprise Information Management** helps organizations fully harness their disparate structured and unstructured data to gain optimal insight from their collective knowledge. We overcome traditional IT barriers to unite all types of files, sources, and even multi-language content in open, sustainable archives. Authorized users of the system can thereby improve day-to-day operations which require access to enterprise knowledge, influence decision-making processes, and gain business intelligence.

ZyLAB Advanced Legal Search Tools

As discussed in George Socha's white paper, all search engines are not created equally. Some, like Web search engines are optimized to display the most popular results. Others, like ZyLAB, are optimized for maximum recall coupled with advanced tools for precision. ZyLAB returns the optimal recall and precision to support a defensible methodology by applying robust search tools, some of which are listed here:

- A complete array of tools to enhance efficiency, such as cross-repository search and retrieval

- Support for large and nested complex Booleans, proximity and quorum search

- Fast fuzzy (supporting first-character changes) and advanced wildcard search (a*, *a, a*a, and *a*)

- Hit highlighting and hit navigation

- Reproducible and reliable relevance ranking

- Forensic indexing of file and document properties

- Automatic language recognition

- Indexing capabilities for compound objects such as nested e-mails, compressed files, e-mail collections, databases, and more

- Extended index and search process auditing and reporting

- Advanced visualization tools

- Incremental indexing of live network data

- Integration with records management, legal hold, identification, collection, legal review, (TIFF) productions and redaction processes

- Advanced text analytics and machine translation

- A search engine mentioned in existing case law.

Choose From eDiscovery Software and Services

ZyLAB delivers high-quality, modular eDiscovery software and services that are tailored to any environment, from serial litigants who are bringing eDiscovery in-house to organizations that need to outsource some of the work for an impending eDiscovery deadline. We offer a full-scale eDiscovery software system behind your firewall and a la carte eDiscovery onCommand services with rapid project turnaround.

ON PREMISE SOFTWARE

ZyLAB eDiscovery & Production System is the field tested, defensible and end-to-end eDiscovery platform preferred by law firms, corporations, the judiciary and legal service providers. By using the ZyLAB technology in-house, many organizations have proven that it is possible to:

- Provide timely responses to litigation deadlines (avoid fines and penalties)

- Provide effective responses (find what is needed, without unknowingly revealing confidential data or non-responsive documents that are not required to be disclosed)

- Save costs on legal-reviews by using automation to organize and process documents for legal review and minimizing the number of documents that must be reviewed manually.

The ZyLAB eDiscovery & Production System is a true eDiscovery and e-disclosure solution that offers our clients all of the most critical capabilities for addressing their eDiscovery and e-disclosure needs. Plus, your IT staff has peace of mind knowing that the deployment is easy to install and maintain, affordable to operate, and compatible with most specialized legal tools.

ON COMMAND SERVICES

ZyLAB eDiscovery onCommand™ is a services-centric solution for organizations that need to immediately respond to an impending eDiscovery and gain practical eDiscovery experience to apply to future cases.

ZyLAB onCommand services provide a full range of outsourced paper and electronic discovery services from preservation, collecting the relevant data, processing, analyzing, reviewing and coding the information, and hosting the data, to production of responsive documents according to your specifications. All services are delivered utilizing our top-rated eDiscovery & Production software.

This onCommand service balances the need for quick turnaround with a built-in upgrade path to on premise deployments.

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

30

31

Sponsored by: **ZyLAB** eDiscovery & Information Management

"In fact, [ZyLAB is] among few companies to address all core phases of the EDRM with our own technology."

"The ZyLAB eDiscovery legal hold software system protects data that is potentially relevant to an investigation from being lost, modified, or overlooked."

ZyLAB Software is Aligned with the EDRM

ZyLAB is a long-term participant in the EDRM community; our eDiscovery & Production System is directly aligned with the 8 core nodes of the EDRM and we integrate with leading presentation tools for the 9th node. In fact, we are among few companies to address all core phases of the EDRM with our own technology. ZyLAB also partners and integrates with leading third parties when applicable.

ZyLAB is an advanced solution that offers the most unique feature set on the market, including:

- Open, modular software construction
- Secure and sustainable long-term data storage
- Best-available search and retrieval capabilities
- A complete, single-vendor solution
- Comprehensive training and full-service customer support

Electronic Discovery Reference Model (EDRM)
© EDRM.NET

The following sections provide an overview of how ZyLAB software fulfills the requirements for each of the core phases of the EDRM.

INFORMATION MANAGEMENT AND IDENTIFICATION

ZyLAB software is uniquely positioned at the convergence of a response to a specific eDiscovery and the implementation of comprehensive compliance and litigation readiness systems. Our software enables you to conduct a thorough early case assessment (ECA) on a data sample or in-place data sources long before other ECA tools and thereby identify potentially relevant data, sources or custodians well in advance. Every engagement benefits from ZyLAB's robust, enterprise-grade information management software.

Features include:

- Fully-featured ZyLAB Information Management Platform
- Early Case Assessment in the wild
- Alerts for files that are scheduled for destruction or retention renewal
- Notifications when certain types of data enters the network
- Auto-classification and foldering of enterprise information
- Automatic bulk conversion to XML and archiving of native format
- Machine-assisted translation of foreign language content

PRESERVATION AND LEGAL HOLD

ZyLAB follows a defensible methodology to copy and preserve the potentially responsive data from any source to a dedicated litigation server and thereby minimize disruptions to operations. Our software copies the pertinent data along with audit details, hash values, and chain of custody records, to a defined location—typically a low cost server or NAS with sufficient disk space to accommodate subsequent processing work. As new data is added to the original source, it, too, is copied on an incremental basis to the preservation archive.

The ZyLAB eDiscovery legal hold software system protects data that is potentially relevant to an investigation from being lost, modified, or overlooked. Our custom, reusable, and traceable legal hold notices and questionnaires can be distributed internally to specify the custodi-

ans and in-network and off-network sources for collection. Features include:

- Auto-wizard to generate reusable electronic legal hold notifications with custom instructions, due dates, auto-populated matter details, and questionnaires
- Bulk merging and distribution of notifications to selected employees in Active Directory® and leveraging meta-data from corporate directories and HR/ ERP/CRM systems
- Legal hold status updates are automatically displayed on the ZyLAB dashboard
- Automatic reminders to non-responsive recipients of the legal hold notification
- Paths to custodians' network data sources (laptop, PST, SharePoint, etc.) are recorded in dashboard
- Batch-copy of data at the defined time (now, day-forward, etc.) from all defined sources to the preservation archive
- Preserves audit information, hash values, and chain of custody data.

COLLECTION AND PROCESSING

ZyLAB eDiscovery software leverages automation and best practices to collect an exhaustive data set and then methodically whittle it down to an optimized review set that is proportional to the matter. We apply our highly-advanced technology to detect and assimilate complex files, cull content, prep it for thorough legal analysis, and continually refresh the collection with new files.
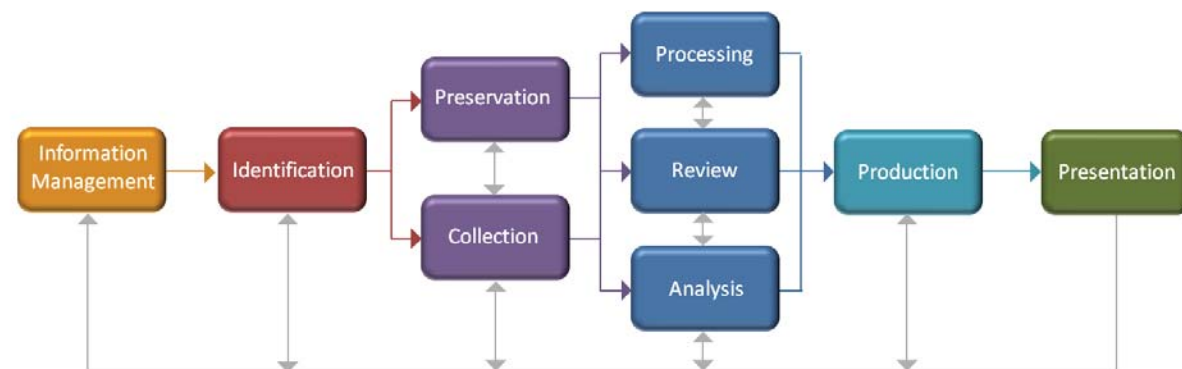
Features include:

- Automatic, incremental collections of new data based on the administrator's defined schedule and frequency
- Inventory reporting
- IT source mapping
- Automatic semantic indexing of the data collection and auto conversion to uniform and searchable format
- Automatic culling to unpack compound files (e.g. zip, rar) and nested e-mails and their attachments
- Support for multiple email formats including eml, nsf, pst, and their attachments
- Integrated multi-directional OCR
- Advanced detection and processing for bitmaps, handwriting, and foreign languages
- Extraction of embedded objects
- Automatic removal of NIST file matches
- Automatic de-duplication
- Automatic language recognition and option for automatic machine translation
- Automatic extraction of metadata
- Automatic recognition and processing of OCR bitmaps
- Automatic flagging of "outlier files" requiring manual attention
- Automatic coding, categorization, foldering and clustering based on defined search engine behavior and results
- Exception workflow.

REVIEW AND ANALYSIS

ZyLAB eDiscovery software helps legal reviewers work more efficiently and ensure there are no surprises in the evidence. The user-friendly, web interface—which can be customized for multi-lingual legal teams—provides simple dashboards and tools to properly move content through the review pipeline. Our advanced legal search tools overcome every review obstacle, including, complex file types, foreign languages, poor scanning quality, aliases, and unforeseen issues in a case.

ZyLAB technology includes the industry's most powerful text mining and semantic search to extract facts, entities, names, code words, synonyms, concepts, and basically, "what you didn't know you didn't know." Review and analysis features include:

- Advanced exploratory search optimized for legal review
- Automatic generation of a robust matter-specific, hosted review portal
- Active hyperlinks to native files
- Concept and pattern extraction
- Text mining
- Metadata enrichment
- Graphical visualization of data
- Content clustering
- E-mail chain & attachment analysis
- Hit highlighting & hit navigation
- Machine translation

- Linguistic analysis
- Privilege logs
- Secured subset review
- Auto tagging, coding and classification
- Keyword sampling
- Annotations
- Customizable interface for international litigation.

PRODUCTION

ZyLAB eDiscovery software accelerates and simplifies the process of preparing evidence for disclosure to opposing parties. The software can add Bates stamps and other unique identifiers to documents, enable condensed printing (multiple pages per printed page), output data into any standard delivery format (e.g., EDRM XML, TIFF, PDF), and compile all of the selected content on a searchable DVD. Likewise, ZyLAB's output files can be accompanied by load files for popular software products like Concordance and Summation. Our industry-leading Intelligent Redaction automates the process of fully removing sensitive data, and our production tools properly mark and output files for third party review.

Defensible Methodology

ZyLAB offers more than just software: ZyLAB offers a full set of proven and well documented methodologies and working instructions to help you defend your approach in

court. This methodology has been developed by ZyLAB in close cooperation with specialist from around the world. Not only can you build a solid chain of custody, you can also show that you have used tools for which extensive case law exists and which have been selected by the most discerning customers in the most demanding environments.

Proven Track Record

ZyLAB has been proven during dozens of history-making cases, including the largest criminal investigation in US history, the largest white collar corporate fraud investigations, and the historic international war crime tribunals. ZyLAB has a documented pedigree that you can trust and reference during your own important cases.

Below are short profiles of just a few of the discovery activities, legal proceedings and organizations that rely on ZyLAB's eDiscovery technologies:

- Investigators and prosecutors in the Enron, Parmalat, Ahold, Worldcom, KPNQuest and numerous other fraud investigations over the past 15 years
- Leading forensic accounting groups such as Deloitte, PricewaterhouseCoopers, KPMG and Ernst & Young
- US Securities and Exchange Commission, US Justice Department, FBI, and OLAF (the European Commission's Anti-Fraud Office)
- The War Crimes Tribunals for the Former Yugoslavia, Rwanda, Cambodia, East-Timor and Sierra Leone

Bringing eDiscovery In-House: RISKS AND REWARDS
by George J. Socha, Jr., Esq.

34

35

Sponsored by: **ZyLAB**®
eDiscovery & Information Management

> "If you are considering bringing all or a portion of the electronic discovery process in-house, the ZyLAB team can provide guidance on what factors to consider."

eDiscovery Consultation

ZyLAB's team of strategic advisors and eDiscovery speci-ficialists are available for introductory or in-depth consulta-tions. If you are considering bringing all or a portion of the electronic discovery process in-house, the ZyLAB team can provide guidance on what factors to consider and how your unique technical environment and information protocol will influence the initiative.

To schedule your complimentary consultation, please call Brad Davis at 1-866-995-2262 ext. 827.

About EDRM.net

The Electronic Discovery Reference Model (EDRM) Project was launched in 2005 to address the lack of standards and guidelines in the electronic discovery market. The completed reference model provides a common, flexible and extensible framework for the development, selection, evaluation and use of electronic discovery products and services. Information about EDRM is available at http://www.edrm.net.

CONTACT INFORMATION:

George J. Socha Jr., Esq.
Socha Consulting LLC
1374 Lincoln Avenue
St. Paul MN 55105
Tel 651.690.1739

Cell 651.336.3940
Fax 651.846.5920
george@sochaconsulting.com
http://www.sochaconsulting.com

Tom Gelbmann
Gelbmann & Associates
290 Grandview Avenue West
Roseville, MN 55113
Tel 651.483.0022
Cell 651.260.5477
Fax 651.483.5938
tom@gelbmann.biz
http://www.gelbmann.biz

Additional Resources

Please contact ZyLAB for more information or for addition-al white papers pertaining to bringing eDiscovery in-house in a controlled and responsible manner. The following resources may also be helpful:

- The Sedona Conference: http://www.thesedonaconfer-ence.org/.
- Andrews, Whit and Logan, Debra and Bace, John (2009). MarketScope for E-Discovery Software Product Vendors. Gartner Research Report, ID Number: G00171281, De-cember 21, 2009.
- Andrews, Whit (2009). Magic Quadrant for Information Access Technology. Gartner Research Report, ID Number: G00169927, September 2, 2009.

- Andrews, Whit and Knox, Rita (2008). Magic Quadrant for Information Access Technology. September 30, 2008. Gartner Research Report, ID Number: G00161178. Gart-ner, Inc.
- Baron, Jason R. (2005). Toward a Federal Benchmarking Standard for Evaluating Information Retrieval Products Used in E-Discovery. Sedona Conference Journal. Vol. 6, 2005.
- Basri, Carol and Mack, Mary (2010). E-Discovery for Cor-porate Counsel, Ed. 2010. Thomson Reuters West.
- Chin, Kenneth (2008). MarketScope for Records Manage-ment. Gartner Research Report, ID Number: G00169707, May 20, 2008.
- Dahlstrom Legal Publishing (2006). The New E-Discovery Rules. Amendments to the Federal Rules of Civil Procedure Addressing Discovery of Electronically Stored Information (effective December 1st, 2006).
- Knox, R. (2008). Content Analytics Supports Many Pur-poses. Gartner Research Report, ID Number: G00154705, January 10, 2008.
- Lange, M.C.S. and Nimsger, K.M. (2004). Electronic Evi-dence and Discovery: What Every Lawyer Should Know. American Bar Association.
- Legal-TREC Research Program: http://trec-legal.umiacs.umd.edu/.
- Logan, Debra and Bace, John (2011). Magic Quadrant for E-Discovery Software. Gartner Research Report, ID Num-ber: G00212221, May 13, 2011.
- Logan, Debra, Bace, John, and Andrews, Whit (2008). MarketScope for E-Discovery Software Product Vendors.

Gartner Research Report, ID Number: G00163258, De-cember 17, 2008.
- Paul, G.L. and Nearon, B.H. (2006). The Discovery Revolu-tion. E-Discovery Amendments to the Federal Rules of Civil Procedure. American Bar Associaton.
- Scholtes, J.C. (2009). Understanding the difference between legal search and Web search: What you should know about search tools you use for e-discovery. Knowl-edge Management World. Best Practices in e-Discovery. January, 2009.
- Scholtes, J.C. (2008c). Maintain Control During eDiscovery. Knowledge Management World. Best Practices in eDiscov-ery, February 2008
- Scholtes, J.C. (2008d). Text Analytics—Essential Compo-nents for High-Performance Enterprise Search. Knowledge Management World. Best Practices in Enterprise Search, May 2008.
- Scholtes, J.C. (2008f). Records Management and e-Discovery: Why we need to re-learn Art of Information Destruction. Knowledge Management World. Best Prac-tices in Records Management and Compliance. November 2008.
- Voorhees, Ellen M. (Editor), Harman, Donna K. (Editor), (2005). TREC: experiment and evaluation in information retrieval. MIT Press.
- ZyLAB podcast series on eDiscovery: www.zyab.com.

Bringing E-Discovery In-House:
RISKS AND REWARDS

ZyLAB®
eDiscovery & Information Management

866-995-2262
zylab.com