The Case for Data and Work Product Reuse

Growth in corporate data stores along with high levels of corporate litigation are driving increases in corporate legal spending. As companies establish their e-discovery processes and tools, they have an opportunity to drive down their legal review and collections costs by making intelligent reuse of data and attorney work product across matters.

Kayvan Vojdani, President at Galaxy Discovery

Presented By:

GALAXY DISCOVERY INC.

GALAXY DISCOVERY INC.

Overview

In today's corporate environment, electronically stored information (ESI) is accumulating at exponential rates creating huge collections of potentially relevant content that corporate legal departments must deal with when responding to e-discovery or regulatory inquires. Relatively high levels of corporate litigation (for many industry segments) combined with continuing data growth are driving more costly, time consuming and complex e-discovery cycles.

To complicate matters, corporate litigation can involve many of the same custodians and require the collection and review of the same content over and over again. For organizations with more active litigation profiles, the same content can be repeatedly collected, reviewed, and produced across numerous unrelated cases driving up the cost of discovery while lengthening the time to respond.

What if a percentage of that previously collected data and attorney work product could be reused across related or similar cases? Should content considered privileged or confidential in one case be considered privileged or confidential in other cases? In many/most cases the answer is, yes.

Lack of attorney work product reuse:

One Fortune 500 corporation maintained an unmanaged and unindexed file share containing approximately **700 TB of files.** Each time the corporation was discovered, that same 700 TB of files had to be searched for relevant content. Several facts were later brought to light with this situation:

- Because this corporation was sued approximately 400 times a year, it was determined the company was collecting and reviewing the same documents and **spending tens of millions of dollars per year** "discovering" the same data over and over again.

- Approximately 42% of the files were so old they were well beyond any retention period that could have been applied and so should have been disposed of years ago.

About this white paper

This white paper explores a variety of challenges found with inefficient and non-standardized e-discovery processes and technologies including the inability to leverage prior work product. The white paper then discusses the potential benefits that can be realized from the adoption of processes and technologies designed to take advantage of prior data collections and attorney work product reuse.

Overview

Data Reuse: The use of already collected electronically stored information (ESI) from one or more legal cases for the benefit/support of another. Data collection in response to discovery can be a long and tedious process involving many internal and external storage repositories (Figure 1).
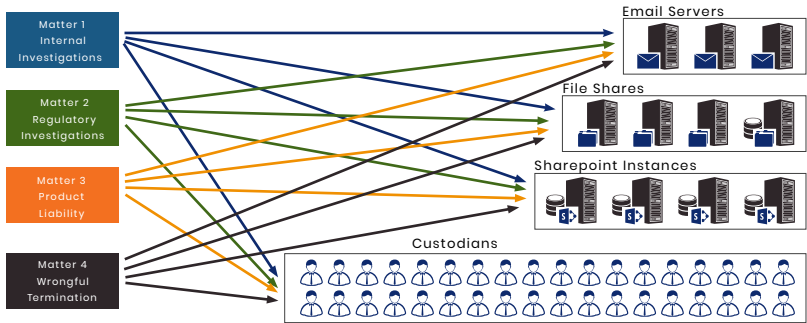

Figure 1: Ineffective information management raises e-discovery cost

Rather than repeatedly searching the many data repositories over and over again, data reuse for discovery takes advantage of prior collection activities by first searching a main cloud-based evidence repository (Figure 2) for previously collected custodian data. Quick access to this already collected data can also enhance your early case assessment (ECA) capabilities by providing relevant content for evaluation early on in the case.
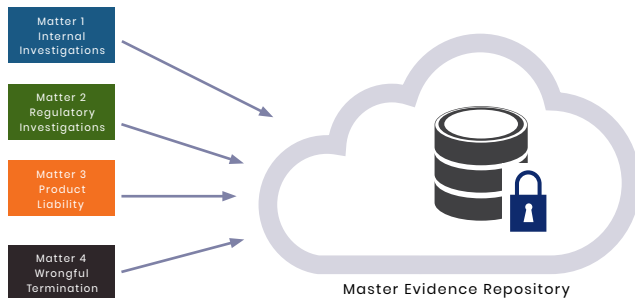

Figure 2:

Data reuse is dependent on centralized evidence repositories

Master Evidence Repository

Related to data reusability is the legal principal of Possession, Custody or Control. Set out in Rule 34(a) of the Federal Rules of Civil Procedure (FRCP) is the requirement that litigants produce documents – including ESI – so long as those documents are in their "possession, custody, or control." This rule makes it clear parties must search, preserve and produce all relevant content - no matter where it's stored – including already existing collected data repositories.

Many corporate legal organizations do not provide a single managed repository for all collected data and instead allow the creation of case specific collection repositories that may not be visible or easily accessible to other non-related cases in the legal department. This situation can cause individual case managers to overlook other non-related e-discovery repositories when conducting a case-specific content search violating the Possession, Custody or Control rule.

When does a data reuse process make sense?
- When cases are related such as class action, product liability, employment, and regulatory investigations
- When common custodians are involved; e.g. "C" level employees
- When common date ranges are involved
- When common or related business activities are involved

Maintaining a single (case-independent) ESI evidence repository for ongoing collected ESI management is both easier and faster to quickly search for ECA purposes than beginning a new enterprise wide collection process. Data reuse does not make sense, however, when the cases, time frames or custodians are less common or there is a high probability that new collection criteria would produce different results.

Attorney Work Product Reuse: The leveraging of already processed (reviewed and tagged) content from one case for use in another. Attorney work product are those written materials, charts, notes of conversations and investigations, electronic files, recordings and other materials directed toward preparation of a case or other legal representation. Attorney work product includes processed content that has been reviewed and determined to be privileged and therefore not subject to review by the other side, or covered by a protection order from a previous case and therefore considered confidential.

Attorney work product also includes content that has been determined to be responsive or nonresponsive in a given case. In specific types of cases such as class action cases, product liability cases, and regulatory actions, content determined to be responsive in one has a high probability of being responsive in a related case.

Utilizing prior content decisions on privilege and confidentiality or duplicative content across cases can provide overall e-discovery savings (Figure 3). In other words, the reuse of coding decisions made on documents for prior related matters can generate measurable time and cost savings when responding to a discovery request.

Figure 3:

Data and attorney work product reuse can save time and money taking advantage of previously coded content

GALAXY DISCOVERY INC.
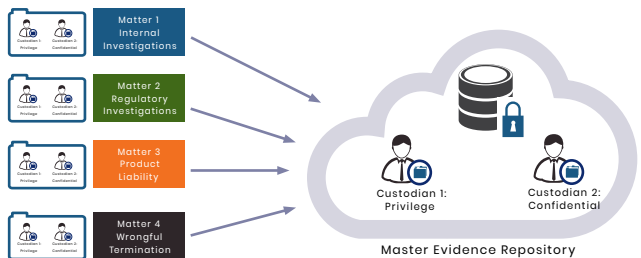
Key e-discovery/regulatory trends highlighting the need for data reusability

Regulatory inquiries are up: The number of annual lawsuits filed against the average organization has been relatively stable over the last 3 years but regulatory actions have increased. In the April 2014 Norton Rose Fulbright Litigation Trends Survey report [1], 52% of respondents reported that over the last 3 years, they had spent more time addressing regulatory investigations than in the past.

> *"Among the largest companies surveyed, 60% have 20 or more lawsuits pending against them, 42% have more than 50."*
>
> Norton Rose Fulbright Litigation Trends Survey report

More data on legal hold for longer periods: Due to several factors including: court congestion; the increasing complexity of cases; and, the sheer amount of data involved in discovery, large amounts of data are being placed on legal hold for longer periods of time, dramatically adding to expanding legal / e-discovery repositories.

Enterprise data stores continue to grow: The ability to quickly research and collect data via the internet, more powerful end-user computer applications, and the seemingly unlimited enterprise storage resources available to employees enable them to quickly collect (hoard) ESI with little oversight or pushback from IT. This mostly uncontrolled accumulation of valueless data was highlighted in the Compliance, Governance and Oversight Counsel (CGOC) survey conducted in 2012 that showed on the average, 1% of organizational data is subject to litigation hold, 5% is subject to regulatory retention requirements and 25% some business value. The rest, 69% of any organization's retained data, has no value to the organization and could be disposed of without legal, regulatory or business consequences. This fact highlights the lack of retention/disposition policy compliance at the average business organization.

Amount of ESI per matter is increasing: Because the amount of enterprise storage continues to expand, the amount of potentially relevant content does as well. Depending on the type of data, the number of pages per GB can vary between 15,000 and 75,000 while the average amount of post culling data per custodian is thought to be between 3 and 10 GB. This means that the average number of pages per custodian per e-discovery can approach 750,000. Multiply that by the average number of custodians in a given case and the total pages that should be reviewed can grow into the millions of pages or more.

Common custodians across multiple matters: In the above mentioned Norton Rose Fulbright Litigation Trends Survey report, the five most active types of corporate litigation in 2013 were labor/employment at 48%, contracts at 36%, personal injury at 25%, IP/Patents at 19%, and regulatory at 19%. All of these litigation types can easily involve common custodians as well as common attorney work product.

Corporate legal departments continue to search for new ways to curb the time required and costs of litigation support including combating the exponential increase in corporate information creation and replication.

E-discovery issues common to most corporate legal departments include:

Time to respond: The time frames for responding to e-discovery requests seem more compressed as compared to years ago due to the fact that there is far more potentially relevant electronic data being collected and reviewed in a given legal action than was the case years ago with mostly paper documents. Many organizations are recognizing that while e-discovery is an event-driven process, by necessity it also needs to be an anticipatory process. An organization must be proactively prepared to respond to an e-discovery request before it arrives by getting their information under control and taking advantage of previous data collections and attorney work product, otherwise they run the risk of an incomplete or late e-discovery response.

Rising e-discovery costs: The cost of enterprise storage continues to fall. Conversely, the total cost to discover content on enterprise storage continues to rise. For example the average price of enterprise class, high availability storage can run in the $4 [2] per GB range while the cost to review 1 GB of content can cost between $18,000 and $30,000 [3]. The cost to review 1 GB of data will no doubt fall over time due to new review automation but that reduction will probably not outrun the expected growth in the amount of data per custodian collected. Even though the cost of enterprise storage continues to fall, the amount of potentially discoverable data continues to rise driving the cost of e-discovery up.

Ineffective ECA: Prior to the "modern era" of electronic discovery, the early case assessment (ECA) process provided attorneys early insight into a case. The seemingly longer timeframes for discovery response, and the relatively small data sets (compared to today) provided the responding legal team time to gather and analyze key information and develop a "go forward" strategy. The move away from mostly hardcopy to mostly electronic data has in turn impacted the nature of ECA. Today, the ECA process is negatively impacted due to much larger data sets, i.e. more documents to and analyze for strategy development. Even with new e-discovery automation, the number of pages of content to find and analyze has grown from thousands or hundreds of thousands to millions or hundreds of millions (and even billions) of pages of potentially responsive content. This dramatic increase in potentially relevant content has changed the ECA process to a mostly data assessment process instead on focusing on finding and analyzing the right relevant documents.

Lack of consistency: A problem many corporate legal departments uncover too late is the lack of consistency in their e-discovery processes. With larger legal departments, individual attorneys tend to create their own specific e-discovery process and decision criteria. Across legal departments these "differences" can create hard to track errors that are much more difficult to document, explain or address versus a standardized process. An often used example of inconsistency is a situation where a document is tagged as privileged or confidential in one case but is not in another.

Benefits of data reuse

The benefits of data reuse for e-discovery include:

Reduced time and cost savings: The reuse of already collected data translates into lower overall costs due to savings from reduced collection and processing costs as well as end-user productivity savings due to less time needed collecting data from end-user equipment.

[2] EnterpriseTech Storage Edition, May 15, 2014: Pure Storage 250 TB All-Flash Array Takes On Disks
[3] InsideCounsel Magazine, May 23, 2012: E-discovery costs: Pay now or pay later

Strategic Insights: Having early access to already collected data enables a faster and more meaningful ECA process to develop a "go-forward" case strategy. For example, early insight into the realities of the case can suggest a strategy of settling as soon as possible, thereby saving resources the company would have spent defending an unwinnable case.

Risk reduction: Both the FRCP and government regulatory investigations impose pressures on the time allowed to respond to data requests. Many organizations have lost cases before they ever got to a courtroom due to an incomplete e-discovery response. The advantage in shortening the length of time to respond to an information request can directly reduce the risk of non-compliance with an information request.

Reuse of prior attorney work product: Without a central collected data repository, the ability to find and utilize prior attorney work product is more difficult and time consuming, if not impossible. A standardized collected data repository will enable attorneys to search across multiple case collections for relevant content as well as take advantage of already coded documents. This next section will outline the benefits of attorney work product reuse.

Benefits of work product reuse

The benefits of work product reuse are dependent on a couple of factors; the amount of privileged and confidential content involved and case similarity. With completely unrelated cases, there will no doubt be very little reuse potential of responsive documents but privileged and confidential content will still have a high probability of reuse. In similar or related cases such as class action, product liability cases, and government regulatory actions, responsive content in one has a higher potential of being responsive in the related case. For cases with attorney work product reuse potential, benefits can be realized from:

Faster e-discovery response time: Calculating the time saved during the data collection phase is difficult because even though you may find some case related data in an existing evidence repository, you still need to go back to all possible content locations to be sure you have found all relevant content for the case. When considering the advantages from an existing evidence repository, data collection cost reduction is less a measureable benefit than the advantage from quick access to relevant data for ECA.

The measureable time savings of work product reuse will be realized during the content review phase by reusing already coded content. For example, let's assume the average legal professional can review and make a determination on privilege, confidentiality or responsiveness on 70 documents of content per hour. Let's further assume that the case includes 1 million reviewable documents. At 70 documents per hour, the time to manually review all 1 million documents would be 14,286 hours or 357 weeks. If 20% of the document collection had already been reviewed and marked privileged and confidential for other cases, the total review-hours saved would equal 2,857 hours or 71 weeks.

The 2012 RAND Report titled: Where the Money Goes points out the biggest cost component in the average discovery process (73 cents of every dollar spent) was the review of documents for relevance, responsiveness, and privilege. Add to that the fact that the amount of electronic data continues to grow driving up the cost of review. To address the rising costs of e-discovery, the reuse of prior attorney work product can measurably reduce the growth of the cost of e-discovery. Continuing with the above example, if the cost of manual review was a conservative $50/ hr., that same 2,857 hours saved would equal $142,857 (Table 1 below).

Time and cost savings calculations

	Assumptions	Costs	Savings
Total number of documents to review	1,000,000		
Review rate - documents/hour	70		
Review cost/hour manual review	$50		
Percentage of work product reuse	20%		
Total review hours		14,286	
Total cost of review		$714,286	
Total review hours saved			2,857
Total review dollars saved			$142,857

Table 1: Cost of internal review with reuse

Higher levels of consistency: For many types of legal actions such as class action suits, product liability, and regulatory investigations, the ability to access and reuse attorney work product from similar or related cases will speed up the e-discovery process but, as important, improve consistency across matters. For example previous coding decisions for privilege and confidentiality as well as prior redactions should not change. An accidental or inadvertent waiver of privilege where content deemed privileged in one or more cases is not tagged as privileged in another can cause the content to become "in play" by opposing counsel. Utilizing prior coding work can greatly reduce this risk of inadvertent disclosure of privileged content.

Attorney work product produced from prior or unrelated cases within the same firm or legal department can be used for educational purposes to understand how/why previous coding decisions were made.

Early strategic insight (ECA): Immediate access to potentially relevant content is key to developing the best case strategy going forward. The availability of a master evidence repository provides the instant access to get started saving valuable time as additional content is collected from the enterprise. This faster ECA capability provides earlier insight so that an informed case strategy can be created.

Architecting for data/work product reuse

Setting up your organization for data and attorney work product reuse is not a trivial matter. Several prerequisites are required to do it efficiently as well as defensively including utilizing common e-discovery tools and repositories, building e-discovery processes that ensure reuse, and employee training.

Acquiring an executive commitment and process owner:

Before any new processes are created and technology purchased, two additional requirements should be addressed;

1. Gain General Counsel level commitment for the new process
2. Assign an owner who will drive the process going forward

If data collection and processing requirements are not correctly (and routinely) followed, cases can be put at risk and e-dscovery costs can skyrocket. Because of this, a high level management commitment is a good idea. To receive a high level commitment, a project manager/process owner should to be chosen early to create and present a business plan (including estimated costs and expected ROI) and a change process plan.

Common e-discovery tools and a central repository:

Many corporate legal departments (especially those spread across different divisions), have not standardized on a common e-discovery tool set, instead acquiring a hodgepodge of differing capabilities. This conglomeration of different tools, which don't necessarily speak to each other, can be difficult and time consuming to use effectively in multi matter situations. To enable the easy reuse of data and attorney work product in e-discovery, the entire organization should be utilizing the same tools and processes.

- Common e-discovery tools should be adopted across the entire organization so that as content is collected and tagged in one case, those same repositories and tags can be leveraged in other cases. As has already been pointed out, content tagged as privileged or confidential in one case should be privileged or confidential in other cases. A common tool set is a key enabler for work product reuse; without common tools, the reuse capability disappears.

- Many e-discovery application vendors now offer on-premise or cloud-based hosted evidence repository services. With a central evidence repository, data already collected for other cases can be included in all case searches to ensure process completeness (see the "possession, custody and control" discussion earlier in the paper) and enable faster and more effective ECA.

Information governance and e-discovery processes:

An efficient and cost effective e-discovery process relies on an organization's ability to actually manage their corporate information. In fact, the Electronic Discovery Reference Model (EDRM) places information governance as the first step in the e-discovery reference model (Figure 4 below). Those organizations without the ability to effectively manage all of their information face the prospect of searching huge amounts of unstructured data repeatedly, much of it expired or valueless, wasting time and money.

GALAXY
DISCOVERY INC.

Electronic Discovery Reference Model

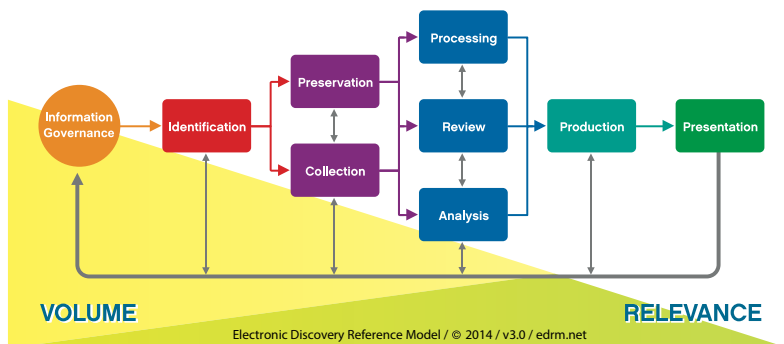Electronic Discovery Reference Model / © 2014 / v3.0 / edrm.net

Figure 4: The Electronic Discovery Reference Model

To position themselves for data and attorney work product reuse, organizations must first understand how they currently handle electronic data to highlight the holes in their IG processes. In reality, if you're not sure how (or if) your ESI is managed, then it's probably not. In addition to fully understanding your current IG processes, the organization's e-discovery processes should be reviewed in relation to the IG process as well.

In the short term, organizations can take advantage of data and work product reuse by creating a reuse policy and processes, designating a reuse program manager, and starting the process of standardizing on a common e-discovery solution set that specifically includes reuse capabilities and centralized collected data repositories.

What to look for in an e-discovery solution

To fully realize the time, cost and risk reduction advantages from effectively reusing data and attorney work product, the right e-discovery tools must be adopted across the entire organization. Capabilities to look for include:

- A centralized collected data repository – as potentially relevant data is collected for different cases, a secure, centralized storage area needs to be utilized so that as different or new cases come into being, this already collected data can be automatically searched as a matter of the e-discovery process.

- Automatic deduplication of collected data to save storage space as well as review time.

- A repository index so that new searches can be conducted quickly.

- The ability to create and manage large numbers of both related and unrelated cases with case specific security measures.

- The ability to create and manage both local and global tags.

- The ability to search on global tag designations across cases.

- The ability to provide full repository reporting.

- The ability to provide conflict alerts so that privileged or confidential content is not inadvertently miss-tagged or released to unauthorized personnel.

Galaxy Discovery enables data and attorney work product reuse

Kayvan Vojdani, the Founder and President of Galaxy Document Services, Inc. and Galaxy Discovery, Inc., has more than 23 years of experience in the legal industry. He started his career at Night-Rider/Ikon Document Services in 1995 and has honed his craft working with leading law firms and national support organizations throughout the USA.

Galaxy is a full-service document-reproduction company that specializes in Advanced Litigation Support. The company has been providing premier support to the legal community in Los Angeles since 1998 to offer superior and affordable outcomes for all document imaging needs.

Galaxy Discovery, Inc., a startup company and new technology division, is a full-service litigation support and e-discovery solution provider. Our services include document scanning, copy services, and e-discovery for specialized law firms, corportate legal departments, and lawyers.
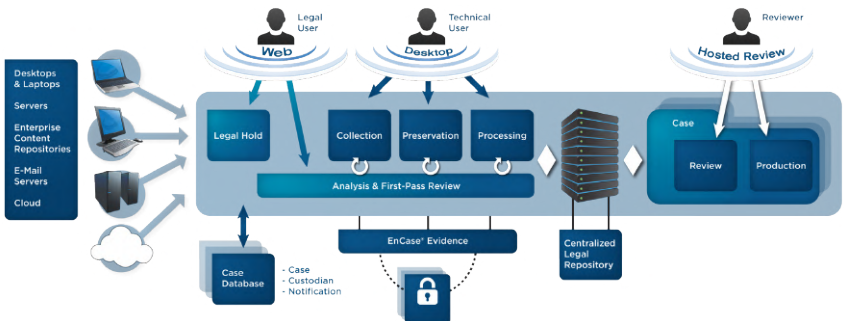
The number of lawsuits companies are dealing with probably won't decrease anytime soon. In addition, the amount of data companies create, receive and store will continue to rise. This combination is the main reason legal budgets will continue to rise. Without a major change in the way companies manage their electronic data, e-discovery will continue to consume large amounts of corporate resources. In the short to medium term, companies can control their e-discovery costs by adopting data and attorney work product reuse processes and technology to both reduce costs as well as risk.

Best practices for data and work product reuse:

- Develop an executive sponsor
- Choose a process manager early in the project
- Create a standard e-discovery process that includes data and attorney work product reuse
  - Evaluate similarities: dates, common custodians, common criteria
  - Concern: is the process repeatable and defensible?
- Standardize on a tool(s) that effectively incorporate data and work product reuse by default
- Train affected employees on the process and tools
- Create a central evidence repository
- Force the use of central data repositories and standardized e-discovery tools
- Establish a process or clearinghouse for all e-discovery data requests
- Apply IG practices to all data repositories (retention/disposition).