

White paper

Best practices for meeting critical eDiscovery challenges



Table of contents

- 3** Introduction
 - 4 Creating a systemized eDiscovery program
- 4** Technology assisted Discovery
- 5** Meaning changes everything
- 5** Advanced eDiscovery and information governance
 - 5 #1 – Email, records, and Consolidated Archiving
 - 5 #2 – Early Case Assessment
 - 5 #3 – Legal Hold
 - 6 #4 – Review and analysis
 - 6 #5 – Processing and production
- 6** Manage, preserve, and secure in place
 - 6 Forensic Collection vs. Forensically Sound Collection
- 6** eDiscovery in the cloud
- 7** Conclusion



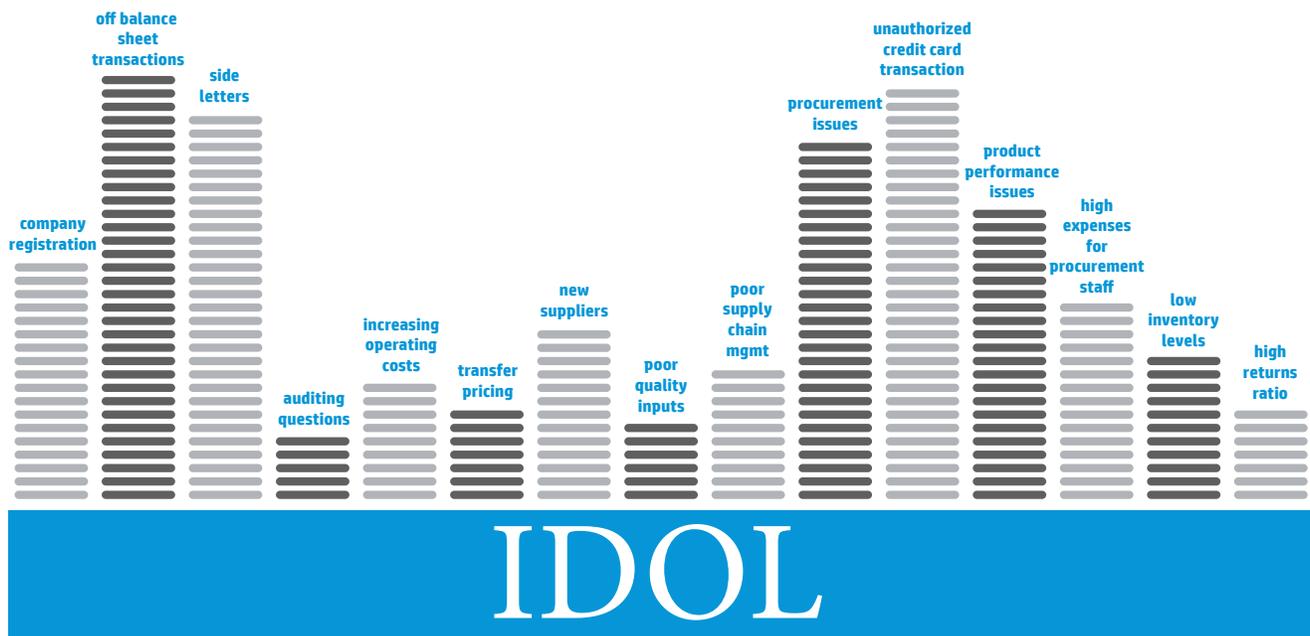
Introduction

Today's organizations are facing sweeping changes in the area of eDiscovery as new and continuing requirements dictate what, when, how, why, and even the type of information that must be preserved and produced. These demands are occurring globally, in step with advancing technology, explosive data growth, the adoption of social media, and pressure to meet evolving legal and regulatory requirements. Mobile access also plays a role, as users connect more and more via networks, the cloud, laptops, and handheld devices such as smartphones and tablets.

At the center of these shifts, organizations are striving to increase visibility into eDiscovery processes and costs—to improve outcomes and streamline IT processes—by not only engaging earlier in the

information lifecycle, but also by leveraging technology assisted discovery to accomplish more in less time and at a lower cost.

A key eDiscovery best practice involves proactive information governance. This straightforward stance is critical because every piece of electronically stored information (ESI) is potentially an asset. But to manage information proactively, organizations must understand what content is in their possession, how it was created, where it resides, who can access it, and if/when it should be deleted. A strategic governance platform can help an organization categorize objects to establish its value, enforce retention and deletion, apply and track legal holds, and finally, enable intelligent search and retrieval for investigations, monitoring, and discovery.



Automatically cluster documents in context by themes.

Creating a systemized eDiscovery program

Implementing a comprehensive governance strategy to anchor end-to-end discovery includes addressing three key areas that can create gaps in the eDiscovery process. These include:

1. Disparate Systems
2. Massive Volumes of Diverse Data
3. Technology

To help address these challenges, three significant shifts in the landscape of eDiscovery technology are occurring. They include:

1. Increased capability to tie many systems together using a platform-based view of information. This capability allows data to be governed by policy while leaving it in place.

2. The capability to interpret information in context. Computer systems can now extract and understand meaning from the vast range of electronic and multichannel content that is generated and stored in human-friendly terms, as well as in databases.

3. Leverage Technology Assisted Discovery, applying smarter technology to assist the legal team in analysis and review of ESI. Technology Assisted Discovery is increasingly important as data volume grows and human friendly information such as email, texts and audio dominate the field.

Technology assisted Discovery

Case law has emerged that addresses issues of discovery burden and proportionality, and the application of Technology Assisted Discovery. For example, in *Da Silva Moore v. Publicis Groupe*, the court finds that the linear review “gold standard” is a myth, that keyword searching is often little more than a game of “Go Fish,” and that Technology Assisted Discovery can yield a fifty-fold savings over manual review.¹

A number of cases discuss the use of technology and automation after laborious, manual review to create a sample set of documents.² The sample set is then used by the computer to find others like it, and put them into a folder for human validation before production. This is generally referred to as automated, or predictive, coding. The broader method of computer assisted discovery includes applying technology to gain insight and meaning early in the matter to help lawyers understand what happened, identify exemplar documents and key custodians, as well as using it during review and afterwards to check coding quality.

A critical component to any discovery process is a reasonable good faith effort and well documented workflow. When conducting paper

¹ *DaSilva Moore*, Case No. 11 Civ. 1279 (ALC) (AJP) (S. Dist. NY 2012) - February 24; June 15 recusal motion denied; gender discrimination case.

² In *Global Aerospace v Landow*, Case no. CL 61040 (cir. Virginia 2012), the producing party gained court approval to use computer assisted discovery over the receiving party’s objections; *Kleen Products, LLC, et al v. Packaging Corporation of America*.

or eyes-on review of every file, one typically follows a conventional workflow that all were familiar with. Now it's reasonable to design a different workflow for different points in the process, for example in ECA using concepts and email communication flows, to align time and money with the amount in controversy. New workflows to leverage Bayesian inference and concept clustering have been in use by lawyers and practitioners for many years.³ At this stage in the evolution of e-discovery some form of computer assisted workflow with machine learning has become a requirement for products that lawyers use.⁴

The impact of the technology shifts described above is significant because they have the potential to dramatically change how legal organizations facilitate governance and eDiscovery in the following ways.

- **Some records can be classified without manual intervention.** For example, communication streams implying insider trading can be automatically captured, regardless of the location of the content.
- **eDiscovery and governance processes can be improved using methods and tools that examine information in context.** Machines have the capability to respond to a query by using the meaning of the content in context with the search request to provide more complete results.

Meaning changes everything

As the pioneer in Meaning Based Computing (MBC), HP Autonomy solutions enable organizations to form an understanding of all information—whether structured, semi-structured or unstructured—and recognize relationships that exist within it. MBC enables computers to ingest and understand the full set of information contained within documents, video files, audio recordings, and social media interactions—not just its associated metadata. This allows computers to harness the richness of human-friendly information, bringing meaning to all data, regardless of what it is or where it resides. Through sophisticated functionality and analytics, MBC automates manual operations in real time to offer true value to the organization.

HP Autonomy's Intelligent Data Operating Layer (IDOL) provides a platform for all Meaning Based Governance (MBG) solutions, transparently providing advanced capabilities, a full range of functionality, and seamless connectivity between all applications. IDOL forms a conceptual and contextual understanding of all content in the enterprise, automatically analyzing information from over 1,000 content formats, 400 repositories, and more than 150 languages. Over 500 operations can be performed on digital content by IDOL, including hyperlinking, agents, summarization, taxonomy generation, clustering, education, profiling, alerting and retrieval.

³ "The Truth About Computer Assisted Review": Concept Clusters Workflow Example, Thomson ARMA CLE, June 26, 2012.

⁴ Gartner in Magic Quadrant for E-Discovery Software, 24 May 2012, also reports, "In 2011, increased emphasis was placed on machine-assisted review techniques, particularly those that use machine learning to help human experts train systems to categorize documents appropriately."

Advanced eDiscovery and information governance

Since the amendments to the Federal Rules of Civil Procedure (FRCP) in 2006 and the ensuing flood of case law, eDiscovery has been an ever-present topic in corporate legal departments and law firms. There are technology tools to help solve the challenge of governing ESI, but organizations need to understand the risks and benefits of using these tools to solve this challenge.⁵

The following five areas comprise the end-to-end eDiscovery and information governance platform:

#1 – Email, records, and Consolidated Archiving

Email and content archiving underpin almost every platform-based governance and discovery strategy. Yet, while messaging systems are particularly well-designed for delivering mail, they are not ideal for high-volume and compliant storage, governance, records management, legal hold, or eDiscovery. Similarly, file shares, SharePoint, and application databases often lack proper governance controls, making them obvious data sources to archive.

HP Autonomy's Consolidated Archive (ACA) provides an intelligent governance layer from which organizations can drive compliance, eDiscovery, records, and information management initiatives, whether on premise or in the cloud and directly from archived data. ACA leverages an innovative split-cell architecture and single instance storage to securely manage distributed sources and attachments.

#2 – Early Case Assessment

Document review is by far the most costly phase of the eDiscovery process. Early Case Assessment (ECA) is becoming a common method for first-pass, high-level analysis and culling to reduce the volume of data sent to review. The deployment of ECA in an enterprise, with its ability to search and analyze data in place, in its native format and within its source repository, is perhaps the most direct example of the mitigating impact of advanced technologies on speed, cost, and risk.

#3 – Legal Hold

Legal Hold as a technology approach has significantly changed in corporate legal departments in recent years. Many vendors initially marketed "legal hold" solutions designed to manage the distribution and tracking legal hold notices, with corporations subsequently perceiving legal hold tools as tactical notification systems. While these are certainly necessary steps, the preservation obligation extends well beyond simple case and notice tracking that requires relevant data to be preserved.

The duty to preserve potentially relevant information can be a disruptive and unpredictable trigger, beginning at the moment when litigation or an investigation is reasonably anticipated.

⁵ See, e.g., K&L Gates, *Case Summaries*, Electronic Discovery Law Blog, available at www.ediscoverylaw.com/articles/case-summaries.

Understandably, this legal responsibility presents the greatest challenge for most organizations. Penalties for the failure to preserve potentially relevant information can include evidentiary sanctions, adverse rulings and fines.

HP Autonomy eDiscovery solutions deliver a single, audited, scalable, and case-based environment to manage all phases of the legal hold process.

#4 – Review and analysis

The biggest cost in eDiscovery is typically the review and analysis phase, which often involves a document-by-document review of data, including both ESI and physical records. While Early Case Assessment provides a method to pre-cull documents prior to review and reduces the data set, the ever-increasing volume of data makes it difficult to control these costs.

While traditional, legacy review technology relies on linear review, which starts with the first document and moves chronologically through the documents, more advanced technology permits non-linear review, with reviewers categorizing and clustering data based on concept and context, thus helping to improve review speeds and allowing reviewers to focus on a given topic.

HP Autonomy eDiscovery offers the most advanced processing, review, and production solution available, and can automatically cluster and visually display related documents, allowing for computer assisted review of all data sources. HP Autonomy eDiscovery enables an efficient, non-linear review and utilizes IDOL's patented concept-based pattern recognition technology for Meaning Based Coding. IDOL learns how reviewers tag documents and automatically suggests coding determinations to save valuable time and improve quality.

#5 – Processing and production

In addition to review, data must be processed to be delivered to the review platform in near-native form, or in the form of load-files, and/or produced for opposing counsel or an investigatory agency. Processing and production tools can provide additional cost savings to counsel by eliminating fees for outsourcing.

While some culling and review platforms can also process and produce data, they often have limited production capabilities, and counsel should be diligent in evaluating the functionality of these platforms. Unlike other solutions, HP Autonomy eDiscovery processes all file types, including instant messages (IMs), audio, video, social media, and web interactions, and imaged documents on the same platform. In addition, HP Autonomy does not rely on to performance optimization techniques such as “jump outs” or partial document indexing. Instead, IDOL's core index engine avoids these techniques with a search capability that complies with the FRCP.

Unique rich and social media processing

HP Autonomy eDiscovery provides the unique ability to perform native audio, video, rich, and social media processing. IDOL's advanced capabilities provide hyperlinking, clustering, automatic

query guidance, and holistic views of audio, video and social media information, enabling legal teams to extract only the information most relevant to a case. Users can cull data sets to the most relevant files and discard those that are not relevant with unprecedented speed and accuracy using IDOL's ability to automatically group data with similar conceptual meaning. HP Autonomy eDiscovery solutions enable organizations to narrow their collections and prioritize files that go on for review.

Manage, preserve, and secure in place

The primary purpose of a collection process is to preserve potentially relevant records. Unfortunately, due to the need to preserve, the collection process itself is often based on a wide net, trapping large volumes of non-relevant materials. To that end, it has become increasingly important to leverage a technique known as “preserve-in-place.” This technique prevents alteration or destruction of the material while analysis is conducted to determine what needs to be collected. With the high costs of collections and processing, having technology that can support a preserve-in-place technique can lead the organization to massive savings in money and time spent on collecting data volumes.

Forensic Collection vs. Forensically Sound Collection

A forensically sound collection is one that is conducted in a forensically sound manner. Forensically sound means that during preservation and collection the file metadata is preserved unaltered (examples of metadata include file name, dates, sender, recipients, etc.). Collecting ESI in a forensically sound manner is required in all legal, regulatory, and governmental document requests and litigation. HP Autonomy Legal Hold offers automated, forensically sound collections via a secure, encrypted connection from desktops, laptops and all enterprise data sources. Collected data is sent to a tamper-proof container with a complete audit trail to manage hash values, file path information, and document authenticity. The ALH Remote Forensics feature is able to search at the sector level, offering protection against lost data with an ability to retrieve and classify information from the “empty space” as well as create full disk images.

eDiscovery in the cloud

More companies are opting for HP Autonomy cloud-based solutions to meet their most stringent privacy and security requirements while implementing a comprehensive information governance strategy directly from the cloud at a level that is unmatched by any other vendor. The shift from on-premise software to on-demand, cloud-based options enables organizations to lower costs and mitigate risk, while defensibly enforcing comprehensive information governance across the enterprise.

Offering the highest level of security available, HP Autonomy data and eDiscovery processing centers offer the following:

- Adherence to Safe Harbor Privacy Principles of notice, choice, onward transfer, security, data integrity, access, and enforcement
- Geographic locations across the globe
- Audited to Statement of Accounting Standard number 70 (SAS 70 Type II)
- Certified to DoD5015.02, SAS 70 Type II, and PCI DSS, UK TNA2002, and Australia's VERS
- Compliance with the U.S.-EU Safe Harbor Framework and the U.S.-Swiss Safe Harbor Framework

Conclusion

In an environment of tightening regulations and continuous litigation, the ability to perform legally defensible and compliant eDiscovery is a concern that every enterprise must address. The volume and complexity of data is now so abundant that organizational processes such as information management, assessment, notification, collection, review, and reporting are too complex and expensive to perform manually. New content types and interaction methods have multiplied with the emergence of social media and mobile devices, further complicating governance and IT processes and directly impacting eDiscovery capabilities.

HP Autonomy is the first and only vendor to offer a comprehensive information governance solution for eDiscovery, Compliance, Archiving, Data Protection, and Enterprise and Legal Information Management all on a single platform. Powered by IDOL, organizations can leverage Meaning Based Computing to index their entire corpus of data and apply a contextual and conceptual framework to all information, structured and unstructured, within documents, video files, audio recordings, and social media interactions—beyond just its metadata. This capability supports every step of the EDRM to ensure eDiscovery that is compliant with the FRCP.

Ranked by IDC as the fastest growing vendor in the enterprise search and discovery sector with market share nearly double that of its nearest competitor, HP Autonomy enables today's corporations, government agencies, and law firms to meet the urgent challenges of information governance and eDiscovery with reliable, proven market-leading technology. With HP Autonomy, organizations can gain full control and visibility over corporate data to confidently react to triggering events from a proactive stance, while meeting requirements for transparency, accountability, and defensibility.

About HP Autonomy

HP Autonomy is a global leader in software that processes human information, or unstructured data, including social media, email, video, audio, text and web pages, etc. HP Autonomy's powerful management and analytic tools for structured information together with its ability to extract meaning in real time from all forms of information, regardless of format, is a powerful tool for companies seeking to get the most out of their data. HP Autonomy's product portfolio helps power companies through enterprise search analytics, business process management and OEM operations. HP Autonomy also offers information governance solutions in areas such as eDiscovery, content management and compliance, as well as marketing solutions that help companies grow revenue, such as web content management, online marketing optimization and rich media management.

Please visit autonomy.com to find out more.





Get connected

hp.com/go/getconnected

Get the insider view on tech trends, alerts, and HP solutions for better business outcomes

autonomy.com

Copyright © 2013 HP Autonomy. All rights reserved. Other trademarks are registered trademarks and the properties of their respective owners.

20130121_RL_WP_HP_Best_Practices_Meeting_Critical_eDiscovery_Challenges

