WHITE PAPER

Addressing Information Governance Challenges

An Osterman Research White Paper

Published March 2014

SPONSORED BY

Osterman Research, Inc.
P.O. Box 1058 • Black Diamond, Washington • 98010-1058 • USA
Tel: +1 253 630 5839 • Fax: +1 253 458 0934 • info@ostermanresearch.com
www.ostermanresearch.com • twitter.com/mosterman

EXECUTIVE SUMMARY

Information governance deals with the management of an organization's information assets – its email, files, voicemails, text messages, social media posts, database content and other structured or unstructured information – in a manner that will allow it to address all of its information requirements and to reduce the risk it faces from improper management of its valuable information resources. Information governance includes the ability to implement litigation holds, respond properly to eDiscovery and compliance obligations, satisfy the information requirements of users across the organization, and use content proactively to gain competitive advantage or improve business operations.

However, good information governance is lacking in most organizations. Most do not retain all of their relevant data as they should; they are becoming inundated by rapid growth in both the quantity and types of information that they must manage; they are losing control of their information assets as they get stored in repositories outside of the organization's control; and they are not sufficiently addressing changes in how information is managed, such as the increasing use of and reliance on mobile devices.

A failure to properly address information governance is serious. It increases the risk that an organization will breach sensitive data or intellectual property, be the recipient of legal sanctions or regulatory fines, or fail to capitalize on the value of the content it possesses.

To minimize these risks, all organizations should develop an information governance strategy and begin evaluating technology providers that can solve its full breadth of governance needs, both now and into the future.

A failure to properly address information governance is serious.

ABOUT THIS WHITE PAPER

This white paper discusses the key problems with current information governance practices and what organizations can do to address the deficiencies in them.

THE GROWING NEED FOR INFORMATION GOVERNANCE

WHAT IS "INFORMATION GOVERNANCE"?

The term "information governance" can convey a wide range of meanings based on the industry in which an organization is involved, the level of regulation or legal scrutiny it faces in the context of its data management, the corporate culture of an organization, and other factors. The following definition provides a holistic view of information governance:

> *"Information governance enables organizations to access and understand human and computer-generated information without bias to repository or location, organize and control this data with a centralized policy engine, and intelligently manage and take action upon this data in accordance with business, legal/compliance, and data management objectives."*

This definition fits well with the thrust of what we discuss in this white paper:

- The requirement to retain important information for the appropriate length of time to meet all organizational objectives, including those focused on eDiscovery, litigation holds, regulatory requirements, improvement of business operations, proper information management and other drivers.

- The breadth of data that the organization manages, including email, files, social media posts, databases and any other structured or unstructured

content that might need to be managed for legal, regulatory or internal business requirements.

- The requirement to retain this information in systems that can preserve it appropriately and make it readily searchable so that it can be produced when needed for all of the relevant audiences within and outside of an organization.

- The ability to delete unimportant information so that corporate risk and cost can be reduced.

WHAT IS DRIVING INFORMATION GOVERNANCE TODAY

There are a number of drivers that are contributing to organizations' interest in information governance:

- **Escalating volumes of information**
  One of the fundamental problems in information governance is the sheer volume of content that must be managed. In many organizations, data volumes are increasing by 50% or more per year, with many organizations experiencing growth in excess of 100% annually. For example, 50% annual information growth will result in data growth of nearly 1,800% at the end of a seven-year retention period. This means, for example, that an organization that today has 50 terabytes of data under storage will have 949 terabytes under storage at the end of seven years. Managing this information properly without appropriate information governance systems in place will be virtually untenable.

- **Proliferating data types**
  The number and diversity of data types continues to expand over time. While email and files have been the primary data types that organizations have had to manage, today organizations must consider management of social media posts, text messages, video content, voicemails and a variety of other information types that may be relevant for purposes of eDiscovery, regulatory compliance or internal use.

- **More places in which data is stored**
  The number of venues in which data is stored has increased significantly over the past several years. Whereas just a few years ago that vast majority of corporate data was stored on internal servers, backup tapes and archiving systems; today data is stored in these venues, but also in employee-managed cloud storage and file sync/share systems, in IT-managed cloud storage repositories, on employee-owned and company-owned smartphones and tablets, on USB drives, on employees' home computers and in other locations.

- **Expanding external requirements**
  eDiscovery, regulatory compliance obligations, legal and regulatory investigations of various types and related types of activities require that organizations be able to govern their information properly or else face a variety of consequences as discussed later in this paper.

- **Expanding mobile requirements**
  A growing proportion of electronic content in most organizations is created and stored on mobile devices. For example, an Osterman Research survey conducted during January 2014 found that 31% of work-related content is created on mobile devices, but only one-half of this content is proactively managed into a central, corporate location. This creates an enormous information governance problem and is motivating many decision makers to pursue governance programs that can address the mobility issue specifically.

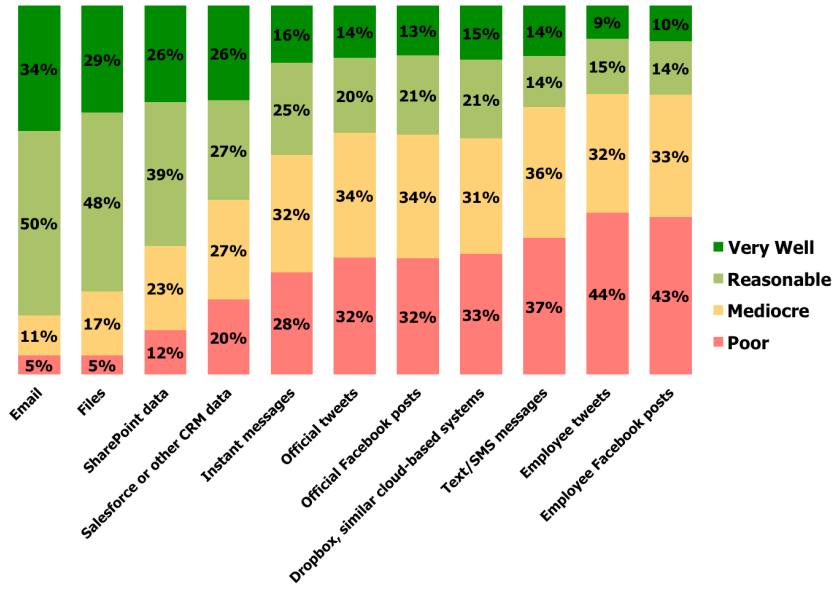
One of the fundamental problems in information governance is the sheer volume of content that must be managed.

ORGANIZATIONS RATE THEMSELVES POORLY

Osterman Research conducted a survey in September 2013 asking organizations the following question: "On a scale of 1 to 10, how well do you think your organization properly governs and manages each of the following data types, where 1 is "we do a terrible job and need to dramatically improve" and 10 is "we govern and manage this data as well as anyone can"?

As shown in Figure 1, we found that for email, files and SharePoint data, most organizations believe they do a reasonably good job at information governance. However, for other data types – which will become increasingly important to govern properly – most organizations view their current governance practices as lacking.

Figure 1
Quality of Information Governance



INFORMATION GOVERNANCE OBJECTIVES

The specific objectives for implementing good information governance vary based on a number of factors, but generally focus on five key areas:

- **Cost savings**
  One of the primary objectives for, and benefits of, good information governance is its ability to reduce corporate costs. On a purely functional level, an information governance program that allows an organization to defensibly delete unnecessary information will reduce the overall storage footprint, resulting in lower storage costs.

  More importantly, good information governance that allows an organization to delete information safely will create potentially enormous costs savings from the reduction of risk. Because unnecessary content may contain information that could be injurious to an organization, eliminating this data also eliminates some level of risk. Moreover, good information governance enables more efficient eDiscovery and the ability to implement legal holds more thoroughly and with the confidence that they will be carried out, further reducing cost.

One of the primary objectives for, and benefits of, good information governance is its ability to reduce corporate costs.

- **Regulatory compliance**
  An organization's records that relate to its business activity are subject to a number of compliance obligations that vary widely by industry and jurisdiction, although it is important to note that all organizations have some level of compliance obligation. These regulations demand the retention of content like financial records, email correspondence between organizations, and employee and client records. Metadata must also be preserved – the Supreme Courts of both Arizona and Washington State, for example, have ruled that metadata must be retained as part of the record of information archived.

  Consequences arising from non-compliance with regulations can be significant and typically involve the imposition of major financial penalties. For example, the Financial Industry Regulatory Authority (FINRA) imposed a $700,000 fine on brokerage firm Piper Jaffray in May 2010 when the firm did not produce 4.3 million emails sent and received between 2002 and 2008. Brian L. Rubin, a member of the law firm Sutherland Asbill & Brennan LLP and former FINRA deputy chief counsel for enforcement, anticipates that FINRA will maintain its attention on brokerage firms' email retention processes and strengthen its examination process of brokerage firms that do not follow up on problems in their retention systems.

  It is important to note that while much of the attention from regulators focuses on email and other forms of communication, files that contain business records – such as advertising literature – are also subject to retention by regulators.

  Many industries experience some form of regulation – although financial services, healthcare, energy and pharma are among the most heavily regulated – and so it is essential to preserve relevant information and establish retention policies that assure compliance obligations are satisfied.

- **Legal compliance**
  An organization currently or about to be involved in litigation has a duty under the Federal Rules of Civil Procedure (FRCP) to preserve all possibly relevant evidence, such as emails, files, databases and other content that could be necessary to have available during the litigation process. This obligation to preserve content normally starts when a party knows, or reasonably should have known, that its data may be relevant to potential litigation. When a litigation hold is required, it is imperative that an organization retain all relevant data, such as all email sent from senior managers to specific individuals or clients, spreadsheets that may contain financial projections, and so forth.

  There are serious consequences resulting from a failure to retain potentially relevant evidence. Courts have discretion to impose a variety of sanctions, such as fines, additional costs for third parties to review or search for data, or even criminal charges. At a minimum, an organization that cannot produce data when required could suffer a damaged corporate reputation.

- **Risk mitigation**
  Another key objective for information governance is the protection of intellectual property that an organization owns or for which it is responsible. Because intellectual property like trade secrets, designs, notes of internal planning discussions and the like are often critical to the future success of an organization, it is imperative that this information be managed in accordance with all legal, regulatory and other requirements.

  Moreover, proper information governance can dramatically reduce the likelihood that sensitive or confidential data will be breached because of poor information management practices. Because a single data breach can

There are serious consequences resulting from a failure to retain potentially relevant evidence.

cost an organization millions of dollars in direct and indirect costs, minimizing the potential for data breaches should be a primary objective of any information governance program.

- **Strategic insight and knowledge management**
  Finally, a key objective of good information governance should be to enable the organization to gain strategic insight about the corporate knowledge that it has stored in electronic repositories of various types. While much is made of "Big Data" initiatives, information governance has as one of its chief benefits the ability for organizations to gather, manage and extract useful intelligence and business value from the store of knowledge it possesses.

POOR INFORMATION GOVERNANCE CAN HAVE SERIOUS CONSEQUENCES

A lack of appropriate governance over the data that an organization possesses can create a variety of consequences, some of them quite serious. For example:

- An organization that cannot produce all required content in a timely fashion in compliance with an eDiscovery order or regulatory audit can face serious fines from courts or regulators, respectively. In the case of *Pension Comm. of Univ. of Montreal Pension Plan v. Banc of Am. Sec., LLC*, 685 F. Supp. 2d 456, 470 (S.D.N.Y. 2010), for example, the court issued an adverse inference sanction because a party acted with gross negligence (as opposed to willfulness) in its failure to preserve electronic documents. The court determined that "contemporary standards" of discovery rendered the failure to preserve and collect electronic files "grossly negligent" and therefore worthy of the severe sanction of an adverse inference, even though there was no proof of intentional misconduct.

- A failure to implement and enforce a complete litigation hold can result in charges of evidence spoliation, resulting in possibly millions of dollars in fines and sanctions. For example, a court found that Samsung, in litigation with Apple, had a duty to impose a legal hold on relevant email starting in August 2010. However, Samsung failed to disable its email system's auto-delete capability and so could not produce relevant content that Apple had requested. This might have resulted in an adverse inference instruction to the jury if the Court had not determined that Apple had also acted badly.

- Poor information governance will increase the likelihood of a data breach that can cost an organization millions of dollars or, possibly, result in the complete dissolution of a business. For example, a failure to implement proper controls over the copying of information to USB drives or other employee-managed media or cloud repositories can result in significant loss of intellectual property or confidential information. Edward Snowden is perhaps the most notable example of one who was able to exploit poor information governance practices, but there have been numerous cases of employees and contractors stealing information in the absence of controls to prevent this from occurring.

- An organization with poor information governance will pay more during eDiscovery because their litigation processes will be less efficient, resulting in the production of excessive amounts of content that must be reviewed by paralegals or attorneys. For example, a RAND Corporation study[i] found that 73% of the costs of producing electronic documents were in review-related activities, while 19% was for processing and only 8% was for collection. As a result, organizations that cull non-relevant documents during the collection and processing phases will incur lower costs. One source estimates the cost of collection at $910 per gigabyte, the cost of processing at $2,931 per gigabyte, and the cost of review at $13,636 per gigabyte[ii]. Based on these

Poor information governance will increase the likelihood of a data breach that can cost an organization millions of dollars.

estimates, every 100 megabytes of content (~1,600 documents)[iii] eliminated during the collection phase will save $1,364 in review costs.

- Employees that are unable to find data because it is siloed and, therefore, unavailable, will spend more time searching for and recreating content that a good information governance system would have made available to them.

- Finally, poor information governance will result in an inefficient expenditure of IT resources. For example, an IT initiative to deploy a better file-sharing solution that does not rely on email as the transport system can provide significant benefits for users who can send larger files, and for IT administrators who now will manage less data on email servers. However, if the file-sharing system does not integrate properly with the existing corporate archiving or data loss prevention systems, there can be an increase in data breach risk, effectively negating the advantages that IT had hoped to achieve.

MOST ORGANIZATIONS HAVE A LOT OF WORK TO DO

DATA TODAY IS SILOED, BUT NEEDS TO BE VIEWED HOLISTICALLY

One of the key problems in starting on an information governance program is that electronic content is heavily siloed. For example, email messages and attachments are stored in an email database, word processing documents and spreadsheets are stored on file shares or in SharePoint, social media content is stored in proprietary content stores, databases often use proprietary file formats and specialized repositories, and various types of electronic content are stored on backup tapes and in data archives.

The compartmentalized nature of content stores means that information governance is more difficult simply because these information siloes cannot share information with one another. For example, a paralegal that needs to find all relevant content that mentions a plaintiff or client must search across a large number of data stores – each with its own interface and method of data access – in order to find all of the content he or she needs. Complicating the issue is that many of these siloes, such as employee-established Dropbox repositories, are largely inaccessible to the organization, especially when this content was managed by employees who are no longer with the firm.

The compart-mentalized nature of content stores means that information governance is more difficult.

FEW ORGANIZATIONS HAVE THE ABILITY TO ADDRESS INFORMATION GOVERNANCE AT SCALE

Another key problem is that few organizations have the ability to manage information governance at the scale they require. For example, many firms simply do not include enough data types in their overall governance solution, perhaps excluding key data types like voicemail or social media content that they might require during eDiscovery or a regulatory audit.

Moreover, the solutions that have been deployed for governance purposes may not be complete. For example, an organization may have a solid email archiving capability in place, but if eDiscovery tools have not been integrated with the archive, the overall process of holding or extracting data during litigation may be inefficient or impossible to satisfy completely.

REALIZING INFORMATION VALUE IS OFTEN OVERLOOKED

Most view information governance as a defensive capability and for good reason: it is useful for demonstrating regulatory compliance, for managing eDiscovery, and for other activities that need proof that an organization has complied with

the law, its corporate policies, legal precedents, etc. Viewed in this context, governance is merely a "necessary evil" to defend against information management problems as they arise.

However, information governance can also be viewed as proactive, enabling decision makers to learn more about their organization, their customers and other aspects of their business. Because organizations possess a vast amount of electronic content, they should view it as a source of strategic insight and business intelligence that can be used to gain competitive advantage, improve employee productivity, speed time-to-market, or offer other benefits.

Key here is that decision makers need to consider information governance as both a defensive tool that can protect the organization, but also as a proactive tool that can provide more insight into operations and business practices.

STAKEHOLDER INTEGRATION NEEDS TO IMPROVE

Another failing of many organizations' information governance practices is that key stakeholders are not adequately consulted and included in the overall process. Because information governance addresses virtually all parts of an organization – senior management, legal, finance, marketing, sales, HR, manufacturing and other functions – key stakeholders from all of these areas must be included in the overall information governance plan. This includes the initial phases of specifying solution requirements, as well as ongoing management of the entire information governance process.

POLICIES ARE INCONSISTENT AND OFTEN MANUAL

Another problem is that policy setting may rely too heavily on manual processes and so can be inconsistent and prone to error. For example, a policy management system that relies on manual updates may fail to apply appropriate policies to all relevant data types, or it may not adequately reflect recent court or regulatory decisions.

MOST DO NOT HAVE THE INTERNAL RESOURCES OR EXPERTISE TO MANAGE THE PROCESS ALONE

Finally, most organizations do not have the internal resources or the expertise to implement and manage information governance on their own. This is not to say that IT departments or others that might take the lead on implementing and managing a governance program could not do so, but rather that they lack the expertise – and often the time or manpower – that are required to manage information governance properly.

Most organizations do not have the internal resources or the expertise to implement and manage information governance on their own.

THE BENEFITS OF GOOD INFORMATION GOVERNANCE

Osterman Research believes that there are five primary benefits of a good information governance plan and execution of that plan:

- **Faster response to eDiscovery requests**
  Responding to eDiscovery requests in a timely manner is critical. FRCP Rule 26(a)(1) requires that all litigants have a solid understanding of their data assets and that they are able to discuss all relevant, data-related issues ahead of the initial pre-trial discovery meeting. Add to this the fact that FRCP Rule 16(b) requires that this meeting take place within 99 days after a legal action begins, and so all parties must have good eDiscovery capabilities in place before litigation starts. In some cases, a court will require even faster production of content. A good information governance program will ensure that eDiscovery requests can be satisfied in a timely manner.

- **Improved compliance with regulatory obligations**
  Robust information governance can also ensure that heavily regulated organizations – especially financial services, healthcare, pharma, food processing and energy firms – can respond to information requests during regulatory actions in a timely manner. For example, organizations that receive a FINRA complaint must respond within 25 days from the time the complaint is served[iv]. Food-service companies that manufacture, import, process, hold or pack food must register with the Food and Drug Administration in order to ensure that threats to the US food supply can be dealt with quickly[v] – the goal is to ensure that food-service companies can respond rapidly to government requests for information after the outbreak of a food-borne illness, for example.

- **More efficient and complete management of information assets**
  Good information governance will also result in much greater efficiency and more thorough management of an organization's information assets. Instead of the common scenario at many firms in which information is siloed, some data stores are not archived or indexed, and some information assets are stored in employees' personal cloud accounts or otherwise out of IT's control, a complete information governance program will ensure that decision makers have access to all relevant information assets using a common interface that provides the assurance that the necessary data can be found quickly and accessed with a minimum of effort.

- **Reduced risk of content spoliation and data loss**
  Another important benefit of good information governance is the ability to minimize – or hopefully completely eliminate – the possibility that data might not be found during an eDiscovery exercise or a regulatory audit. Evidence spoliation during a legal action, for example, can carry with it enormous and damaging consequences. For example, an adverse inference instruction from the court to a jury, in which the jury is instructed it may assume the party failing to present data is hiding something, can be extremely damaging. In the case of *United States v. Suarez* [2010 U.S. Dist. LEXIS 112097 (D.N.J. Oct. 19, 2010)], the court allowed the jury to infer that the plaintiff's deleted text messages were favorable to the defendants in the case[vi].

- **Improved employee productivity**
  Finally, good information governance will enable employees to be more productive because they will spend less time searching for content and less time recreating the content they cannot find. If we assume that the typical user could save only five minutes per day searching for content because of their access to a good information governance system, it would result in a savings of nearly 21 hours per employee per year.

Good information governance will enable employees to be more productive.

RECOMMENDATIONS

HOW TO GET STARTED

Getting started with an information governance program involves three key steps:

- First, implement an appropriate archiving solution that will ensure the retention of all information types that the organization will need to retain. This includes the obvious types of content like email and files, but also social media posts generated from company-authorized accounts, text messages sent from company-owned mobile devices, voicemails, content in cloud-based storage repositories, log data, etc. In many cases, work-related content from employee-owned devices, social media accounts, and cloud-based data repositories will also need to be retained.

- Next, legacy data cleanup should be implemented in order to eliminate superfluous or outdated structured and unstructured content that may be safely deleted. Determination of specific data to be deleted must be performed on a case-by-case basis, but this is a critical step both to reduce the cost of information governance and to reduce risk.

- Organizations should also implement appropriate records management and electronic content management solutions that will ensure the appropriate management of corporate data and make it available to all stakeholders in a timely and complete way.

TAKE A HOLISTIC APPROACH

Decision makers must realize that they do not have a "data" problem or a "structured vs. unstructured data" problem. Instead, they have a problem with information and how it is managed. This is an important distinction because of the strategic nature of information management in the context of addressing not only an organization's eDiscovery or regulatory requirements from a defensive perspective, but also from the perspective of how the organization will use information to gain competitive advantage and improve its operations. Consequently, decision makers must view their information governance activities holistically and ensure that management of information is viewed from this strategic viewpoint.

DON'T ASSUME INFORMATION GOVERNANCE HAS ALREADY BEEN ADDRESSED

Many decision makers mistakenly assume that they either have addressed their information governance needs by simply implementing a backup solution, or they believe that managing just a small proportion of the information they generate will address all of their information governance needs. Instead, decision makers must look at the big picture across all of their data types and information management processes to determine if they are addressing actual information governance. This will include not only addressing legal and compliance requirements, but also proactive information management that will satisfy the requirements of every stakeholder across the organization.

VIEW IT AS AN ONGOING PROCESS

Information governance may begin as a small project in one operation of an enterprise, but it must be viewed as an ongoing process, not simply a series of standalone projects. This viewpoint is essential if organizations are to develop and maintain a robust and active information governance system that will meet all of their information management requirements moving forward.

INCLUDE ALL RELEVANT STAKEHOLDERS

Another key element of a good information governance program is to include all of the relevant stakeholders that have an interest in seeing that corporate information is managed properly. In an Osterman Research survey conducted in September 2013, we found that IT was considered an information governance stakeholder in 96% of the organizations surveyed, followed by legal in 87% of organizations. However, from there fewer stakeholders were included: 78% considered HR a stakeholder, followed by compliance (59%), risk management (45%), finance (25%) and marketing (11%).

A failure to include all relevant stakeholders in the information governance process will result in a solution that falls short of meeting corporate expectations. For example, a solution that is specified only by IT and legal will probably address eDiscovery issues well, but may fall short of meeting regulatory compliance obligations. Moreover, if finance, marketing and other functions are excluded from the information governance discussion, it is highly unlikely that the solution that gets implemented will be able to address Big Data requirements

Information governance may begin as a small project in one operation of an enterprise, but it must be viewed as an ongoing process.

and other, more proactive uses of the solution.

FOCUS ON CONTENT DELETION TO REDUCE RISK

One of the fundamental benefits of an information governance solution is its ability to reduce legal and regulatory risk arising from a failure to find all relevant content. However, it should also be viewed as a means of reducing the risk associated with retaining unnecessary data for too long, the risk that comes from users missing important content that is relevant to their work, or the risk of missing opportunities that might otherwise be obtainable if analytics could be applied to corporate information.

IMPLEMENT THE APPROPRIATE TECHNOLOGIES

It is essential that any information governance solution be based on the appropriate technologies that will ensure that all relevant information is captured and managed appropriately. The solution should allow management of corporate policies and information in a highly scalable way to accommodate the enormous growth of information and expansion in data types that most organizations are experiencing. The best practice for implementing any information governance system is to opt for the solution set that offers management of all data types that are required today and will be required in the future, but that is sufficiently modular so as to permit a phased introduction of the solution at a pace that meets corporate requirements.

SUMMARY

Information governance – the ability to intelligently manage and take action upon human and computer-generated data in accordance with business, legal/compliance, and data management objectives– is an essential best practice, but one that many organizations are not doing well. Although many organizations have implemented ”governance” solutions, they often do not govern all of the information that should be retained, nor do they manage their information assets in a holistic way that will benefit all parts of the organization.

To address this problem, an information governance solution should be implemented in virtually every organization so that legal, regulatory and other information management needs can be addressed. This solution should be able to retain, manage and delete all relevant information types in a way that will address every functional requirement in the organization.

One of the fundamental benefits of an information governance solution is its ability to reduce legal and regulatory risk.

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, web pages, etc. 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. Autonomy’s product portfolio helps power companies through enterprise search analytics, business process management and OEM operations. 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.

REFERENCES

[i] *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery*, RAND Institute for Civil Justice

[ii] http://www.slideshare.net/jmancini77/arma-michigan

[iii] http://www.providusgroup.com/doc_review/doc_review_calc.php

[iv] https://www.finra.org/web/groups/industry/@ip/@enf/@adj/documents/industry/p006746.pdf

[v] http://www.seafoodsource.com/es/seafood-and-international-trade-law/23582-food-facility-registration-mandatory-with-the-fda

[vi] http://www.ediscoverylaw.com/2010/11/articles/case-summaries/court-imposes-adverse-inference-for-failure-to-preserve-text-messages-related-to-criminal-investigation/