Information Governance comes of age

How to take back control of your enterprise data
Table of contents

3  Taking back control of enterprise data
4  Why govern information?
4  Why traditional Information Governance approaches fall short
5  A contemporary approach to Information Governance
6  HP Autonomy’s Information Governance portfolio
7  Getting started
8  Conclusion
8  About HP Autonomy
8  To learn more
Once considered a problem that was too complex and expensive to solve, Information Governance is experiencing a rebirth. As escalating volumes and greater requirements on data are fueling demand, new solutions are emerging that allow organizations to balance short-term needs with a long-term governance strategy.

**Taking back control of enterprise data**

The need to govern enterprise data is not new. Virtually every organization has too much data to manage and has long aspired to place better controls on this information. But there are three forces at work that are now driving Information Governance from a back room problem to a hot topic in boardroom discussions:

1. **The information challenge is getting worse** – Not only is the volume of enterprise data doubling every 12-18 months, but it also is now being stored and accessed almost everywhere. An increasingly mobile workforce using multiple devices to access data is forcing organizations to react and deliver better access—from all places and at all times. This, combined with increased regulations and requirements on data, is escalating the problem and driving up risk and cost.

2. **We’re discovering the value of enterprise data** – Traditionally, Information Governance has been perceived as a measure to control risk and cost. While there are clear benefits to this end, recent industry discussions about Big Data have helped raise awareness about the benefits of enterprise data as well—for example, business intelligence, knowledge center access, and enhanced productivity. With the added focus on potential top-line revenue growth, the need for Information Governance is at an all-time high.

3. **Technology is now catching up with demand** – Until recently, the answer to Information Governance was a massive and unwieldy platform that required deep integration and a high level of manual work (to classify and migrate data), which resulted in unclear ROI. Much has changed in the recent past. Today, solutions are emerging that are much more modular, allowing organizations to solve a specific governance problem first (with proof of ROI) and then add on additional modules to achieve even greater value. This approach, combined with elevated capabilities to automate retention policies from a centralized dashboard, is quickly removing historical barriers.

This white paper delves deeper into the above topics and also suggests some simple, cost-effective approaches to getting started with your own Information Governance strategy.
Why govern information?

There are many reasons to better control and manage information. The business benefits that are most commonly associated with Information Governance include:

- **Lower risk** – By gaining better control of information, enterprises are able to guard against accidental data leaks and ensure greater compliance with data management regulations (internal or externally mandated). Perhaps most important, organizations with proper governance controls can better protect themselves from inadvertent data spoliation. This added measure to ensure data is preserved when it’s required for legal matters or investigations can mitigate risk of sanctions, fines, or even adverse inference decisions in court that can lead to damaging lawsuits and decreased shareholder value.

- **Lower cost** – Despite user desires to keep all data forever, this practice simply isn’t cost effective or practical. According to Gartner Research, enterprise data is projected to grow 40 to 60 percent—with unstructured data projected at a growth rate of up to 80 percent1—resulting in large volumes of invaluable or duplicate data. Some estimates show that as much as 69%2 of enterprise storage is made up of files that are of little value to the business. A better approach is to identify the truly important data that should be maintained, and to defensibly dispose of the rest. This yields a significantly smaller storage footprint that ultimately translates to lower IT and administration spend, as well as more efficient legal holds and cost-effective eDiscovery. Placing better controls on enterprise data also allows IT departments to more efficiently tier storage or migrate data to the cloud (with hygiene) for better cost containment.

- **Better productivity** – Maximizing the use of information and reducing search and knowledge identification times increases business efficiency. According to recent findings by IDC, the average knowledge worker spends about 20%3 of their time searching for information. This drain on productivity can be reduced by better unifying information silos and providing access to more full-featured search capabilities across a broader set of data.

- **IT efficiency** – In some cases, technology can simply be bogged down by the amount of inactive data that it must access. By relocating the portion of this inactive data to an active repository such as an archiving or records management product, IT departments can dramatically improve operational efficiency and protect themselves from certain SLA violations.

In short, forward-looking enterprises can achieve a wide range of benefits with an Information Governance strategy. Steps can be taken to achieve one or two benefits right away with a point solution, or they can build toward and end-to-end solution with a portfolio of modular technologies.

Why traditional Information Governance approaches fall short

While it’s relatively easy to define Information Governance objectives broadly, organizations have historically faced several roadblocks when it comes to realizing the benefits. This is due to the inherent flaws in existing alternatives. Two alternatives are:

1. **A massive migration plan.** Because few Information Governance solutions can stretch across data silos, information has often been moved to a centralized location where controls are then applied to all data now residing in this single repository. The exercise of moving and duplicating data to centralize management in one system imposes a cost and complexity that many companies simply cannot afford. It also introduces significant risk of data loss and spoliation, as migrations of this size can often incur hidden exceptions. Even when moved, it’s is difficult for the receiving system to understand the “compliance state” or meaning of file types such as video, audio, social media, and email. As a result, additional custom tools and workflows are often created to supplement the central system, and these often create even further risk and complexity.

2. **Band together a string of technologies that are not optimized to work together, or implement a platform that attempts to do the same.** Such a strategy often results in a massive index that creates additional storage requirements and costs, is difficult to manage, prone to risk, and a challenge to audit for completeness.
In both cases, organizations often find they are dissatisfied with the breadth of enterprise data that they can govern. Frequently, there are gaps in data types or repositories that can be accessed, which can lead to sub-optimal results when applying legal holds, searching, and performing eDiscovery, among other actions. As an example, many organizations are increasingly looking to govern both unstructured data (human-generated information like email that is difficult for a computer to understand without additional context) and structured data (computer-generated information that falls into neat rows and columns that can easily be sorted) in a single Information Governance model. Addressing this combination of data often falls apart with the two strategies above because the data is very different and difficult to reconcile.

A contemporary approach to Information Governance

An alternative and more-modern approach to Information Governance starts with the recognition of these key tenants:

• Almost any data, in almost any repository, written in almost any language is a target for governance and thus must be effectively accessed and understood
• A centralized policy layer that will help automate (what were previously) manual actions is required to organize and control data efficiently
• Moving data to a repository is not always feasible or cost effective, making the ability to administer some data in place is paramount

Equally important, an Information Governance solution must be comprised of a breadth of technologies that can be implemented individually to solve a specific problem, or be combined with relative ease to deliver a complete end-to-end solution. The former allows organizations to take an early step into governance and perhaps build a holistic Information Governance roadmap, while also building a business case for the point solution they have implemented. The latter will allow an enterprise to achieve the full breath of benefits outlined previously.

And finally, a modern approach to Information Governance recognizes that there are many constituents within the organization that the initiative must serve. This requires technologies that enable action to be taken upon governed data, in accordance with varying business, legal/compliance, and data management objectives.
HP Autonomy’s Information Governance portfolio

HP Autonomy addresses the call for a contemporary approach to Information Governance by delivering a portfolio of modular solutions that help organizations access and understand human- and computer-generated information. Our approach understands information in its context—without bias to repository or location—and organizes and controls this data with a centralized policy engine, so that you can intelligently manage and take action upon data in accordance with the varying needs of unique user across your organization.

This portfolio includes four key market offerings:

• **Information Archiving** – Helps organizations centralize data from key information sources (for example, email, IM, social media, voice and video, plus structured data), across the world and across the enterprise. Delivered as a cloud-based or enterprise-class software solution, the world’s largest archiving provider (based on total data archived) helps customers proactively prepare for litigation, comply with wide-ranging industry regulations, and dramatically lower storage costs.

• **Enterprise Content Management** – Helps organizations address today’s critical barriers to enterprise productivity and information governance: extreme data volumes, multiple data repositories, and disruptive forces such as a cloud and mobility. As a result, organizations gain a full understanding of information across data silos, enable more productive collaboration, comply with regulations, prepare for legal and regulatory inquiries, protect and secure information, and ensure data retention and disposition.

• **eDiscovery** – Helps corporations, law firms, and government entities prepare for, and react to, legal matters and investigations involving large amounts of electronic data—regardless of data type or where it resides. A pioneer in technology-assisted review, HP Autonomy offerings span the entire Electronic Discovery Reference Model (EDRM) framework—combining Early Case Assessment, processing, and review in a single offering and allowing customers to avoid the risk and cost of switching tools or vendors.

• **Data Protection** – Enables enterprises to efficiently protect a new generation of Big Data spread across a range of locations, applications, and formats. These solutions work seamlessly with HP storage products to deliver comprehensive, easy-to-use backup and recovery solutions that maximize storage efficiency and performance, enabling customers to fulfill regulatory compliance, archiving, and litigation requirements with speed, flexibility, and cost-effectiveness.

All four market offerings are powered by HP Autonomy’s IDOL technology, which ties them together and delivers significant benefits. IDOL provides access to hundreds of data types, reaching across myriad data silos and throughout the enterprise. It also relies on HP Autonomy’s patented conceptual technology to understand the meaning of human-generated content, thereby delivering automated classification and policy application of unstructured data. Further, because IDOL can connect to hundreds of different information silos, it uniquely enables the ability to administer data in place, providing organizations with a choice of how to govern their data. This powerful combination sets HP Autonomy apart from alternative approaches to Information Governance and ultimately delivers tangible benefits to customers.
Getting started

Not all organizations are ready to implement an end-to-end Information Governance solution. Some simply want to address a specific problem, prove the ROI of that solution, and then add on to achieve greater benefits over time. HP Autonomy has several solutions for those organizations as well.

One alternative is to start with an information archiving or records management system that captures and manages data from the point of implementation going forward. But for many enterprises a bigger problem is legacy and inactive data that was created in the past. This data—both structured and unstructured—can be very expensive to store and can pose significant organizational risk. By applying technology to better control this data, you can ensure that it is stored in a central place and managed with uniform policies and rigor. Starting here also gives you important experience in establishing and testing policies before moving ahead with broader deployments like records management or archiving.

HP Autonomy has two solutions to address the legacy data problem: one for unstructured data and one for structured data:

• **HP Autonomy’s ControlPoint** helps you gain access to, understand, classify, and reduce outdated, unstructured legacy information. This unique technology helps organizations shine a light on “dark data” that is sitting unmanaged in email repositories, file shares and SharePoint sites—ultimately allowing for defensible data disposition or the migration of targeted legacy data to information archiving or records management systems. This results in more efficient legal holds and improved access to information.

• **HP Autonomy’s Application Information Optimizer (AIO)** helps you gain access to, understand, classify, and reduce outdated structured legacy information. This technology relocates inactive data from production systems and legacy databases while preserving data integrity and access. In doing so, it allows organizations to retire outdated applications and optimize the efficiency of production databases—and ultimately migrate a portion of this formerly inactive data to an active repository such as an archiving or records management system.

Both ControlPoint and AIO deliver short-term cost savings (among other benefits) by reducing an organization’s storage footprint and can be implemented specifically to clean up outdated and invaluable information that is no longer needed. These products can also be used to gain a snapshot of what data is contained within a specified repository so that you can deliver critical insight that informs a broader governance roadmap for organizations early in the Information Governance maturity model. And again, for organizations farther along in the maturity model, both technologies can migrate targeted data to a long-term/active repository so historical data can be managed in unison with go-forward data as part of an end-to-end solution.
Conclusion

Information Governance is no longer an aspirational objective or a “boil the ocean” strategy, which it was when the term was first in vogue a few years ago. New technologies allow you to access, understand, control, and take action on a wide variety of data—across data silos and across the enterprise—to address business, legal and compliance, and IT objectives. Equally important, these solutions are delivered modularly, so an enterprise can start with one distinct goal and then add on to achieve greater benefits. With these advances, and an increasing need for such solutions, now is unmistakably a coming of age for Information Governance.

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. 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.

To learn more

To learn more about Information Governance and HP Autonomy’s portfolio of modular solutions, please visit: http://www.autonomy.com/informationgovernance

© Copyright 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.