

Business white paper

# Turn your information into enterprise value

HP Autonomy's Policy-Based Information Governance solution



# Table of contents

- 3** Governing in the era of information
- 3** Why govern information?
- 4** Governing information for visibility, status, and policy
- 5** How traditional content management approaches fall short
- 5** A new approach: policy-based information governance
- 7** Increase your effectiveness at a lower cost
- 7** Implementing policy-based information management
- 8** Comprehensive, meaning-based information governance
- 8** About HP Autonomy

## Govern information

### Saving everything: good approach or costly habit?

While some organizations feel it is more cost effective to simply save everything, believing that low cost disk storage has little impact on costs, throwing storage at an information overload issue can actually drive up costs in several areas, including:

- Servers and infrastructure
- Cooling and heating
- People to manage infrastructures and information
- Unchecked growth from duplicates
- Backup expenses
- Data conversion to ensure data is readable over time

As a business leader, you are constantly adapting to manage ever-increasing volumes of data stored in disparate systems. In this new era of information, it is all too clear that businesses must proactively manage the retention and management of information for regulatory compliance, business value, and IT efficiency. Using information governance (IG) practices, it is possible to put in place formal methods that will enable you to unify people, processes, and technology to achieve your business goals and obligations.

Our highly regulated and litigious environment provides numerous reasons to govern information accurately and consistently. However, without a process for governing how information is managed, retained, and ultimately destroyed, your organization is sure to uncover areas of legal and financial risk. Much of this risk can be caused by duplicate files stored across various systems.

When risk is present, the warning signs are clear: reduced employee productivity, spiraling storage and eDiscovery costs, and the issues of meeting regulatory, privacy, and compliance requirements. All of these challenges can result in reduced competitiveness and profitability. However, with the right policies defined and enforced, you can control the amount of information under management, and defensibly dispose of “obsolete” information.

Given the broad distribution of information across multiple systems, enforcing governance system-by-system is not a practical approach. Today, more than ever, information crosses corporate boundaries with greater frequency. For this reason, governing information based on what crosses the firewall is not practical. A smarter approach involves developing governance policies for individual information assets, regardless of format or location. In this way, you can enforce policies regardless of where the information is consumed, enabling you to effectively govern information across your enterprise.

## Why govern information?

The need for information governance is driven by three key factors. If achieved, these capabilities can help you increase efficiency, streamline access to information, and meet compliance obligations more effectively.

• **A reduction in storage footprint translates into greater IT savings.** As new forms of information including email, video, audio recordings, images, and social media proliferate in the mainstream, enterprise information volumes have skyrocketed. The amount of unstructured content globally grows at a rate of approximately 80 to 90 percent a year, much faster than information stored in databases. SharePoint and other productivity tools create additional silos, which leads to duplicate content. Some estimates show that as much as one-third of information stored is made up of duplicate files that are of little value to the organization. In the case of enterprises moving to the cloud for greater efficiency, storing duplicate data can take away from the potential efficiencies to be gained through these new deployment types.

• **Maximizing the use of information and reducing search and discovery costs increases business efficiency.** According to recent findings by IDC, the average knowledge worker spends from 15 to 35 percent of their time searching for information.<sup>1</sup> This drain on productivity can be reduced with effective information governance processes that increase efficiencies in search and knowledge management.

• **Increasing the effectiveness of regulatory compliance programs and processes reduces business, legal, and regulatory risk.** The past decade has brought a dramatic increase in regulatory and industry compliance initiatives, each with requirements to manage, retain, and produce information based on specific requirements. Privacy laws such as HIPAA and PCI impose restrictions on how information is stored and who may access it. In the courts, regulators are unwilling to tolerate delays or accept the excuse that preparing information for a request comes at the cost of compliance. These factors and others are driving businesses of all types and sizes to reconsider how they manage information to ensure they can meet these mandates. Monetary penalties, as well as the impact on a company's brand, are high and continually increasing.

<sup>1</sup> <http://www.kmworld.com/Articles/Editorial/Features/The-high-cost-of-not-finding-information-9534.aspx>

Despite the opportunity for significant business impact, most organizations are overwhelmed by the volume, scope, and complexity of the information governance challenge. Numerous studies have shown that while some organizations have a documented information management strategy, it often goes unreferenced. Taking this approach, however, is risky, and can leave your business open to a wide range of consequences.

## Reining in the high cost of eDiscovery

Managing information for compliance or to ensure litigation preparedness is more complicated when content is hard to find, walled off in silos, or “hidden” in formats that key programs cannot decipher. Particularly in an environment of growing volumes of diverse information, the need to review enormous amounts of content can drive eDiscovery costs into the hundreds of thousands, if not millions. These costs are increased further by the need to use systems that require data to be moved or copied, a process that can create huge numbers of duplicate files. However, when information is managed using solutions that leverage meaning based computing, you have the ability to understand the content and context of information, enabling you to remain prepared to find and leverage the right information when you need it.

## Governing information for visibility, status, and policy

Gartner defines information governance as “the specification of decision rights and an accountability framework to encourage desirable behavior in the valuation, creation, storage, use, archival, and deletion of information. It includes the processes, roles, standards, and metrics that ensure the effective, efficient use of information in enabling an organization to achieve its goals.” To meet these criteria, information governance best practices must include three main capabilities, as follows:

- **Visibility into the location and nature of all information assets.** With information stored in hundreds or thousands of locations, the first step in implementing a solid information governance program is to be able to location and access all content across the enterprise. For many organizations, information is stored across hundreds if not thousands of locations and there is no comprehensive catalog or taxonomy to identify where it is. For these reasons, finding information in a typical enterprise requires systems that can search, identify, and interrogate content across the full spectrum of enterprise systems.
- **An understanding of the compliance “status” of all information.** Policies can only be applied to information if you understand its business and compliance value, for instance, by knowing the following:
  - What is the business purpose of the information?
  - Is the content governed by a legal hold or current investigation?
  - Is it a record, or relevant to a current project?
  - Does the information contain sensitive content that requires compliance with Personally Identifiable Information (PII) laws, PCI Data Security Standards, the HIPAA Privacy Rule, or other regulations?
  - How long since the information was last accessed?
  - How is the information secured and who has access to it?

Knowing the answers to these questions can help you decide how and where to store the information, how to govern it, and when and how it should be disposed. But this can be particularly difficult to achieve for some organizations because most unstructured information is not easily classified in terms of typical metadata (filename, date last stored, author, etc.). Gaining an understanding of the compliance status of information requires you to understand its meaning, regardless of whether it is audio or video, social content, or an email.

- **The ability to classify and act on information according to information governance policies.** Once the value, or risk profile, of an information asset is understood, you need a way to classify or apply the proper information governance policy. With policies applied, you can take action on the information and automate certain steps, such as assigning a retention policy. This allows you to manage, archive, and retain information according to specific policies or respond to investigations or inquiries such as a FOIA or audit requests. When information is categorized to meet policy and compliance guidelines, you have the ability to perform a range of cost- and time-saving actions to streamline information governance.

A few examples include:

- Moving information to a lower cost storage based on its organizational “value”
- Eliminating duplicates
- Classifying information according to records policy
- Applying policy to all information assets

- Encrypting high-risk assets for greater security
- Applying security Access Control Lists (ACLs)
- Disposing of content after a required retention period
- Implementing a legal hold to prevent spoliation or disposition

The ability to perform solid information is impacted by the sheer volume of content—across emails, video, audio, images, and other files. And it can be nearly impossible applying policy to content using a manual system. For this reason, you need a way to automatically apply policy to vast amounts of content across disparate systems, based on its compliance status.

## How traditional content management approaches fall short

While it is straightforward to define what is needed for effective governance, there are also several roadblocks that can get in the way of implementing governance processes. Because most systems today are limited to managing content in their own repositories, information is often moved to a centralized location where additional technology systems are used to govern information. The cost of moving and duplicating data to a centralized management system for operational or compliance purposes imposes a cost and complexity that many companies find cost-prohibitive. These approaches are often used because current systems are unable to understand the compliance status or meaning of file types such as video, audio, social media, and email. When this happens, businesses often begin adding custom tools and capabilities that create added risk and complexity.

Unfortunately, the fallback position is to simply do nothing. However, deciding to not implement a governance program can have serious consequences for your business, including the following:

- In legal terms, the lack of action is an action, thereby creating liability
- Avoiding establishing an IG program locks a company into a never ending cost/risk spiral
- Indefensible IG practices can result in legal action, fines/penalties, or loss of reputation due to a legal prosecution (whether it is successful or not)
- Ongoing IG inefficiencies require moving/copying data and can drive up storage costs
- Allocating IT budget to unnecessarily high storage costs reduces capital available for innovation and strategic projects

## A new approach: policy-based information governance

Running your business using solid information governance practices requires the use of systems that enable you to understand the meaning of information in context with other content across the enterprise. HP Autonomy's policy-based information governance solution provides a next generation information platform that delivers a range of capabilities that allow you to:

- Automate governance processes
- Proactively mitigate risk
- Ensure legal preparedness
- Drive growth through enhanced productivity and collaboration
- Limit the movement and duplication of information

Implemented with HP Autonomy's IDOL (Information Data Operating Layer), HP Autonomy's Policy-based Information Governance solution helps you automate the integration of people, processes, information, and technology using a unified, secure, and scalable platform that understands the meaning and concepts held within information. These capabilities are performed independent of file type, repository, or language, helping you govern distributed and diverse enterprise information at a lower cost.

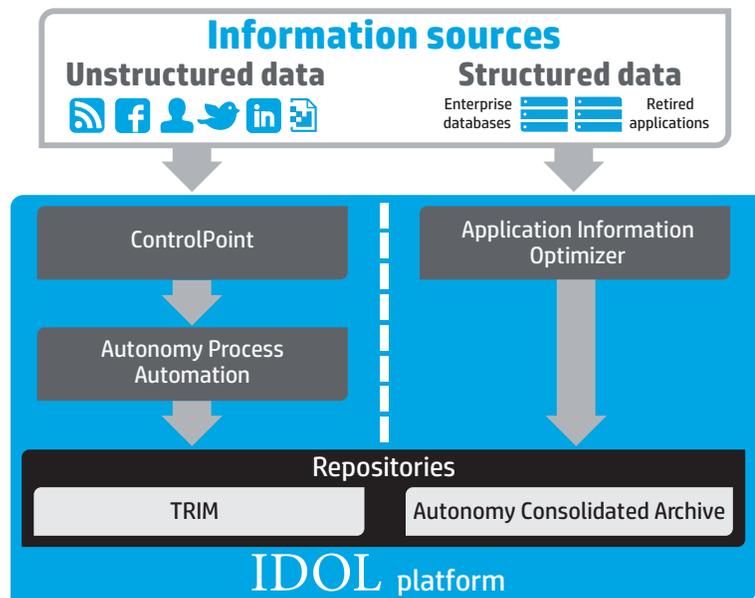
HP Autonomy's Policy-Based Information Governance solution consists of three core products that operate together, and may be deployed with other HP Autonomy solutions to address broader information governance needs.

- **HP Autonomy ControlPoint** is a centralized policy hub for analyzing the current compliance status of enterprise content. ControlPoint applies compliance policies to distributed content across repositories such as Documentum, Filenet, OpenText, SharePoint, file shares, email systems, and internal applications. ControlPoint enables you to inventory legacy data, and eliminate redundant, outdated, and irrelevant content.
- **HP TRIM** helps legal, government, and enterprise organizations to effectively control corporate records by enforcing retention and disposition policies. TRIM simplifies the capture, lifecycle management, security, and disposition of all corporate records.
- **HP Autonomy Process Automation (APA)** is a full-featured, intelligent process automation engine that offers easy-to-use visual process modeling tools, rules-based routing, intelligent electronic forms, and real-time business activity monitoring, to move content to archive and records systems, manage disposition workflows, and support other compliance processes.

The core capabilities of HP Autonomy's Policy-Based Information Management solution may be extended using the following Autonomy products:

- **HP Autonomy Application Information Optimizer (AIO)** enables you to archive legacy data from existing applications or retire the complete data store from existing applications to reduce application footprint and improve performance. AIO offers a rich set of capabilities required to control and manage database growth and apply governance best practices.
- **HP Autonomy Content Archive (ACA)** provides a scalable, unified content archive that controls and manages electronically stored information and assists with litigation preparedness. ACA enables you to understand information in any format, repository, or language, allowing you to perform advanced analytics and draw intelligence from information in real time.
- **HP Autonomy eDiscovery** offers a complete suite of applications for early case assessment, preservation, rapid investigation, and automated review. The eDiscovery solution delivers an innovative approach by enhancing productivity and reducing costs through configurable review workflows and dynamic Meaning Based Coding.

HP Autonomy's Policy-Based Information Governance solution helps you identify, classify, and take action on all content across the enterprise, based on regulatory policy, as shown below.



Autonomy Policy Based Information Management is comprised of a set of integrated products that work together to enable enterprise wide governance of virtually all information

## Increase your effectiveness at a lower cost

The advantages of using HP Autonomy's Policy-Based Information Governance can have a significant and positive impact on your business. The solution can help you cut costs in a number of areas, automate processes, and develop a new level of trust in your organization. The solution enables you to:

- **Reduce spiraling infrastructure and storage costs** by limiting duplication, managing hierarchical storage, and enabling defensible disposition of outdated or non-business-related materials.
- **Reduce eDiscovery costs** by decreasing the volume of content and duplicates, which can result in less time spent reviewing and classifying information in response to a litigation, audit, or regulatory inquiry.
- **Reduce the risk of fines, penalties, and loss of reputation** that can result from regulatory breaches or premature destruction of content.
- **Automate manual tasks** to reduce inconsistencies, enhance productivity, and minimize the overall cost of governance.
- **Enhance records management** by enabling the automated classification of records, and making it easier to search all records, as needed.
- **Create a new level of trust in the information produced in the organization** by implementing systems that ensure accuracy, consistency, and brand integrity.

## Implementing policy-based information management

Because information governance is a process and not a single action, it is important to set expectations and gain an understanding of where to start. Knowledge of what information exists, where it is located, and its current status is the first step in developing an informed approach to information governance.

An important part of the overall solution, Autonomy Control Point allows you to survey and identify the location and current compliance status of all information assets. This provides an assessment of the magnitude of the issues that exist within your organization and gives you information to create the basis of your plan for moving forward.

With a more solid understanding of the scope of the governance challenge, you can prioritize based on a "highest value/lowest risk" status. Following are some of the common steps that organizations take as they begin building an information governance program:

- **Legacy content reconciliation and big data** – After years of acquisitions, many organizations end up with legacy repositories and no idea how to govern them. Even organizations that have not grown significantly through acquisitions are sitting on terabytes, if not petabytes, of information. Legacy "modernization" and "big data" projects all require data cleanup to be cost effective, making sure that information with no value to the business is not moved from one infrastructure to another. With HP Autonomy's Policy-Based Information Governance solution, you can clean up legacy data stores, for instance, prior to a cloud or infrastructure migration.
- **Policy-based, federated records management** – As you know, records and content can be created and stored across many repositories. This makes it essential to govern records consistently based on type rather than native application or source repository. It is also key that users have a seamless experience with continued access via their familiar applications. Using a federated approach to records management allows you to meet these important goals by allowing information to be managed in place, allowing users to stick with their familiar applications for access.
- **Classification and governance of email, SharePoint files, and other content** – The need for email archiving is well understood, but archiving all email creates unnecessary complexity and adds to storage costs. Appropriately classifying content in each repository, with appropriate processes in place for de-duplicating and archiving according to business value and purpose, will help you to enforce cost-effective information governance.

## Comprehensive, meaning-based information governance

HP Autonomy's meaning-based approach to information governance offers you a wide range of capabilities that span the enterprise, allowing you to meet your needs for managing information throughout the information life cycle.

### Meaning Based Information Governance

A single platform approach for effectively managing information, enterprise-wide, in accordance with corporate/organizational policy and business goals

**Gain full understanding** of all information including its location, business meaning, and governance status

**Automate policy implementation** across all repositories, data types, and social media interactions in a consistent, effective manner

**Prepare for and respond** to audits, investigations, and eDiscovery and compliance obligations to mitigate risk and reduce cost

**Protect, archive and secure** information to meet business SLA based on its meaning and value

**Ensure retention and disposition** of records and non-records in a defensible manner in accordance with organizational policy

## Summary

HP Autonomy's Policy-Based Information Governance solution can help you transform enterprise information into value by taking a meaning-based approach to managing information. The solution is comprised of an advanced set of technologies that enable you to:

- Manage documents, email, records, and virtually all other content types residing in numerous repositories
- Perform eDiscovery to meet the requirements of the Electronic Discovery Reference Model
- Archive using an advanced solution that allows you to retain information according to business value
- Preserve information using meaning-based data protection that empowers organizations to manage the explosive growth of data with advanced backup and recovery capabilities—from the edge of the network to the datacenter, and across physical, virtual, and cloud environments.

## 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.

Please visit [autonomy.com](http://autonomy.com) to find out more.

Sign up for updates  
[hp.com/go/getupdated](http://hp.com/go/getupdated)

