

Brochure

# Gain control over all enterprise content

HP ControlPoint



## A better way to manage big data

Most organizations today store data in a number of business systems and information repositories. This practice has taken hold over the past two decades with the rapid advancements in information-generating technology. But as the volume of data continues to increase, organizations are facing significant business risk and loss of efficiency because they cannot effectively control the information contained in these siloed repositories—and sometimes are not fully aware of what information they have.

The unprecedented growth of data is mostly due to the increasing numbers of business and personal computing platforms, such as Microsoft Office applications and SharePoint, cloud computing, mobile devices, tablets, social networks, email, and line-of-business applications. If not properly governed, SharePoint is a prime example of how a collaborative business system can contribute to big data, and potentially create dark data within information silos across the enterprise.

Understanding the value of data in active enterprise systems and legacy systems enables you to better manage your information footprint and associated storage costs through defensible destruction of duplicate, obsolete, and trivial data. Identifying legacy data and establishing its status as sensitive or secure information enables you to minimize the risk of accidental misuse and leakage.

## Identify, connect, and control data across enterprise systems

HP solutions help you to gain control over all information residing in systems across your enterprise. HP ControlPoint and HP IDOL allow you to identify and connect your information siloes, and apply policy to this data to ensure its appropriate and ongoing management.

### HP ControlPoint

HP ControlPoint leverages HP's Intelligent Data Operating Layer (IDOL) and connector framework to identify, analyze, and control content across enterprise repositories. ControlPoint enables classification, categorization, and policy application to content indexed by IDOL. ControlPoint's executive dashboard provides valuable business insight into information; themes, locations and value.

As a centralized information governance console for all connected data sources, ControlPoint simplifies the definition and application of policy regardless of data format or location. There is no need for source-specific policies, which become difficult to manage and unify across the enterprise, and you eliminate the need for staff to learn multiple system tools and user interfaces. ControlPoint's role-based security allows the delegation of different tasks to reduce the likelihood of errors and bottlenecks.

**Figure 1:** ControlPoint reporting dashboard with 2D cluster map showing hot zones.



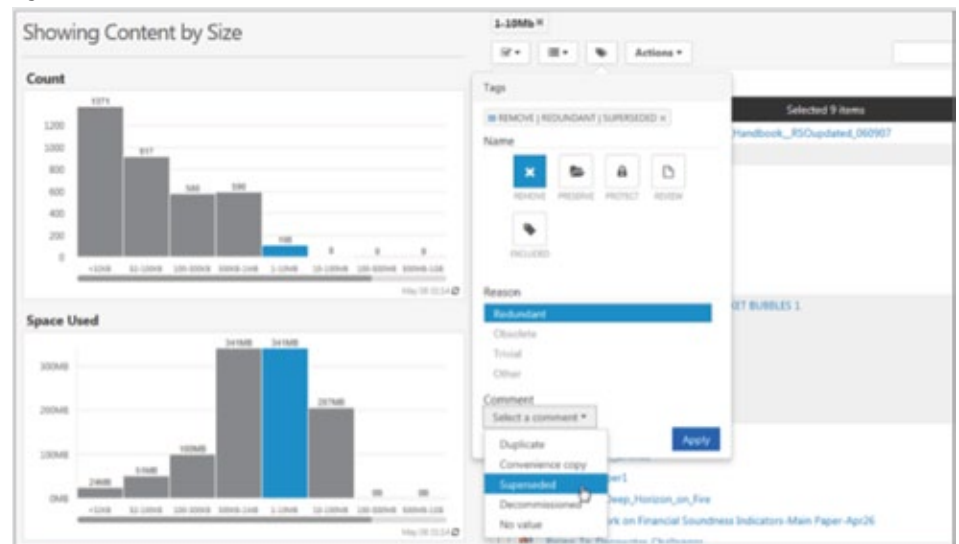
ControlPoint provides advanced, graphical visualization of information clusters based on meaning. This graphical representation makes it very easy for executives to identify trends or popular concepts and themes. Two-dimensional cluster maps show heat zones of information grouped by concepts, and clicking on the graph displays the underlying documents within that zone. Three-dimensional spectrographs display how these clusters change over a period of time. Clusters offer valuable understanding in the evolution of your enterprise information and are useful in training policies to apply defensible disposition management.

ControlPoint’s executive dashboard presents the analysis of indexed data graphically with statistical summaries and data categorization. This analysis allows you to understand the breakdown of data types and categories across the enterprise or specific repositories. A range of tools and reports support a policy-driven and compliant clean-up process, with the following capabilities:

- Reports may be generated on items marked for deletion,
- Maintain audit trails of policy selection criteria and execution,
- Review and approve workflow processes

The ControlPoint dashboard serves as a collaborative hub, bringing together risk managers, compliance officers, and legal counsel to collaborate and enforce governance policies in one system. Navigational tools allow you to drill down to explore different areas of policy module application. For instance, an officer can learn how much data is potentially on hold across all enterprise systems, or what potential policy violations are occurring. Comprehensive reporting brings critical transparency to your company’s compliance condition.

**Figure 2:** ControlPoint dashboard with statistical summaries



### HP IDOL

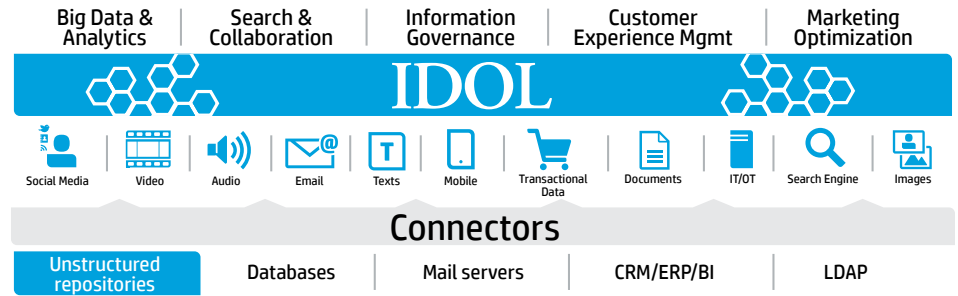
HP IDOL works at the core of ControlPoint to index all information, making it visible, transparent, and available to be analyzed, actioned, controlled, and governed. Information sources such as file shares, SharePoint, and Microsoft Exchange are indexed through the IDOL connector framework. IDOL forms a conceptual and contextual understanding of the content, giving you the ability to index and analyze information from over 1,000 different content formats. IDOL’s mature connector framework allows users to search across the entire enterprise using a web interface, providing an unprecedented view of all information assets and enabling searches that are fully compliant with FRCP requirements. IDOL also supports education or entity extraction such as finding personally identifiable information (PII) in documents.

### Connector framework

Through IDOL’s connector framework, ControlPoint can access and analyze data from a range of sources. Standard connectors provided with ControlPoint include:

- File shares
- Microsoft Exchange
- Microsoft SharePoint
- HP TRIM

**Figure 3:** IDOL connectors make information accessible across devices and applications



## Get more from your data and reduce footprint, risk, and cost

Information contained in multiple systems and repositories provides little insight into business performance and themes if it is not managed in a holistic manner. Siloed information provides tactical value to a small segment of the overall enterprise, but rarely delivers strategic value to executives or the enterprise as a whole. The ability to report on and display an organization’s information evolution is invaluable for future strategy and performance.

### The emergence of dark data

Dark data is the data that you may not know you have. This currently unmanaged, often unknown electronic information resides in various repositories across the organization. Dark data is mostly human-readable, unstructured, unindexed, inactive, and orphaned. Proliferation of dark data can result from the “Bring Your Own Device” (BYOD) phenomenon, along with the continuing explosion of big data with new unstructured data types such as audio, video, and social media. In this scenario, information governance challenges arise when information that is generated and stored on mobile devices, social platforms, file-sharing services, and SharePoint goes unmanaged. Because so little is known about this dark data, it places the organization at risk.

### Identify redundant, obsolete and trivial data

The information footprint of an organization can be significantly reduced to provide measurable returns, if redundant, obsolete, and trivial data can be identified within system repositories.

Redundant data consists of duplicates in the form of unauthorized copies of documents, emails, records, or database information residing in file shares, SharePoint sites, mail systems, and databases.

Obsolete data consists of information that is no longer in use, or is out of date. When determining whether data is obsolete, you can identify its creation date, last modified date, or access date and then assess this information in conjunction with an appropriate retention policy.

Trivial data is determined by file type, where the file type has no content value, such as executables, system files, and thumbnails.

**Figure 4:** ControlPoint dashboard showing the ratio of redundant, obsolete and trivial data.



### Legacy data clean-up

Organizations often look to legacy data cleanup to address a number of information governance issues. The process may be undertaken as a standalone activity to address a specific problem such as part of an onboarding process for applications and information acquired during merger and acquisition activities. Legacy data clean-up may also be undertaken as part of an ongoing business project to improve efficiency and reduce information footprint. Another use is to conduct legacy data clean-up to gain greater insight into data holdings and the categories they fall into, aiding in the establishment or improvement of information governance practices.

ControlPoint's Legacy Data Clean-up component provides opportunities for analysis and data tagging in legacy repositories to ensure policy is applied and appropriate action is taken to ensure defensible disposition or ongoing management.

Analysis is conducted on a range of data attributes including date fields, file properties (type and size), creator, category matching, custom fields, and duplicate assessment against defined masters. ControlPoint's Legacy Data Clean-up process consists of five stages moving from identification through to action and clean-up.

Stage	Process
<b>Identify and index</b>	<ul style="list-style-type: none"> <li>Identify data sources</li> <li>- Index at meta data level (redundant, obsolete, and trivial)</li> <li>- Index at content level (deeper analysis, insight, data relevance)</li> </ul>
<b>Analyze</b>	<ul style="list-style-type: none"> <li>Statistical summaries (% of redundant, obsolete, and trivial)</li> <li>- Data creation and volumes over time</li> <li>- Categorization of data based on content index (concepts and context)</li> <li>- Personally Identifiable Information (education)</li> </ul>
<b>Organize</b>	<ul style="list-style-type: none"> <li>Identify information clusters and common content patterns</li> <li>- Data groupings and categories</li> <li>- Apply policy according to categorization and classification</li> </ul>
<b>Reduce</b>	<ul style="list-style-type: none"> <li>Policy-driven defensible disposal</li> <li>- Audit logs and reports to support compliant processes</li> </ul>
<b>Manage and migrate</b>	<ul style="list-style-type: none"> <li>Data identified to have business value and relevance migrated and managed under the organization's Information Governance plan</li> <li>- Declaring documents into the records management system</li> <li>- Secure moves and holds</li> <li>- Migration to tiered storage or repositories</li> <li>- Ongoing application of policy to new data based on established and trained categories</li> </ul>

## Ensure information governance and control across enterprise systems

### Information silos and non-compliance

The risk of non-compliance can be high for organizations living with information silos or SharePoint site proliferation, which can easily lead to the existence of legacy data and dark data. Approaching each challenge separately is costly and leads to a disjointed information governance implementation that is likely to have gaps.

Through HP IDOL's unified and consolidated view and connectors, ControlPoint can apply standardized policies from a central policy engine. The data is effectively managed through a central information governance console.

ControlPoint offers policy-driven classification, auto-declaration, manage-in place capabilities and seamless integration with HP TRIM records management software to help you bring information under control within a robust information governance framework.

### Reduce risk and meet the demands of regulation

Meeting internal governance and regulatory compliance requirements and responding to legal discovery, external investigation, and audit inquiries are both business obligations and challenges. Non-compliance is a risk that may result in financial penalties, interruption to your business operations, and negative publicity. To meet compliance obligations and be prepared for legal discovery or investigation, organizations should adopt a proactive approach to information management. When you work proactively, it is much easier to foresee areas of risk and address these before they become real problems.



### **Records management & defensible disposition**

Many organizations keep their content well beyond its expiration date or inadvertently delete valuable business documents. These information management practices may result from disconnected information silos, a lack of understanding about what content exists and where, poor policy application and enforcement, and a lack of information security. ControlPoint allows you to manage or dispose of content in place for subscribing systems or declare content in these systems as records and move in to HP TRIM for their ongoing, secure and compliant management.

### **Seamless integration with HP TRIM electronic document and records management software**

ControlPoint's seamless integration to HP TRIM electronic document and records management software facilitates robust, compliant records management for all declared business records. ControlPoint's auto-declaration of records simplifies the flow of content from enterprise systems into HP TRIM. HP TRIM manages, secures, and provides access to all corporate records in business context, ensuring authenticity, integrity, and reliability.

ControlPoint with HP TRIM provides full document and records management functionality, automating the retention and disposition of information. ControlPoint provides a range of audit logs and reports out of the box, allowing you to keep a defensible history of policy changes as well as the effects of policy application to managed content. These user-driven or automated reports help you monitor the application and effect of policies on your information. HP TRIM manages records according to applied disposition schedules and event triggers, ensuring secure management and defensible disposition.

## **Increase efficiency and elevate productivity through automation**

As the volume of electronic content continues to grow, it is critical for businesses to adopt an information governance solution that does not rely exclusively on the manual efforts of users and administrators. ControlPoint effectively automates the consistent application of policy to content based on the conceptual understanding of information in a multitude of file formats. Additionally, you can use document/location metadata as part of policy assignment. ControlPoint also enables you to automate information policy enforcement and audits.

### **Train ControlPoint categories**

The ability to train ControlPoint categories based on existing IDOL categories or records within HP TRIM removes much of the burden of having users manually create or map categories. A selection of documents from system repositories is used for training and benchmarking, ensuring categories are based on meaningful concepts and real business context. This capability improves the efficiency and accuracy of categories and the application of policy to content.

Draft categories can be prepared without affecting documents in production systems. These can be refined and tested to determine the impact they are likely to have on enterprise documents. A category may be refined by adjusting the weighting of a term, the selection threshold, or by adding field text. This can be done individually or in combination. Once refinement is complete, the category can be published, making it available for use in automatic policy execution against content managed by ControlPoint.

### **Automate categorization**

Categorization is critical to the application of policies. Whereas traditional collaboration and ECM systems rely upon users to categorize and tag information on an individual basis, ControlPoint leverages IDOL to analyze information, arranging it into self-similar groups or clusters and leveraging IDOL by matching the data against trained categories. Through duplicate and near-duplicate identification, storage costs are minimized, while documents identified as records can be declared into the records management system (HP TRIM). Users can file content using a single click or content can be automatically classified based on rules and IDOL's understanding of the concepts and context contained in the documents.

### **Automate policy-driven classification**

Once data is categorized, policies can be applied to the data for ongoing management. Policies can be created with keywords, metadata, and/or example documents using a simple Web-based, wizard rich dashboard. Policy creation is intuitive and allows automatic enforcement. Policies can be linked to the classification scheme within HP TRIM, ensuring the appropriate retention and disposition schedules are applied to declared records.

ControlPoint enables you to:

- Automate policy application governing all aspects of the information lifecycle, including deletion prevention, storage management, and ultimately disposition management by applying meaning-based policies at data creation
- Leverage a web-based, non-IT centric dashboard to create policies
- De-duplicate data across repositories to minimize storage costs and reduce discovery times

### **Manage-in-place**

ControlPoint provides the flexibility for users to perform specific actions on content in-place. These in-place capabilities simplify the management of enterprise content according to business value and lifespan. The ControlPoint client makes it easy to identify and categorize content from any subscribing system, then based on policy, this content can be moved into a records management system (HP TRIM), the archive, or indicated as ready for destruction.

ControlPoint leverages IDOL indexes and categories to apply policy to information via the connector framework. The policy can dictate a number of actions to be carried out, including:

- Hold
- Release hold
- Copy
- Secure copy
- Move (between repositories)
- Apply tags
- Delete
- Declare record
- Initiate an HP Process Automation (APA) workflow

Content that has not been accessed for a defined period of time can be moved to more cost-efficient storage, eliminating the need to license costly SQL servers in SharePoint environments (inactive SharePoint sites) and alleviating the strain on network resources. Both of these capabilities can bring substantial savings to your organization.

Email that falls under an automatic clean-up rule can be checked to see if it matches a records category before deletion, minimizing the risk of important information being deleted inadvertently if the user hasn't actively declared it as a record. The efficient management of these information platforms throughout their lifecycle greatly reduces storage and infrastructure costs.

## **Summary of key benefits**

- Gain visibility and understanding of business risk associated with information stored in enterprise systems.
- Automate compliance, legal hold, and disposition management based on policy and the conceptual understanding of information across various file formats.
- Perform legacy data cleanup and defensible disposition by :
  - Identifying, categorizing, and classifying content in enterprise information silos
- Inventory content based on business value and lifespan to facilitate appropriate management
- Set a policy to drive secure management of business records in the records management system or to defensibly dispose of redundant, outdated, and trivial content

## 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)**



Share with colleagues

---

© 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.

