

IBM Unstructured Data Identification and Management

Discover, recognize, and act on unstructured data in-place



Highlights

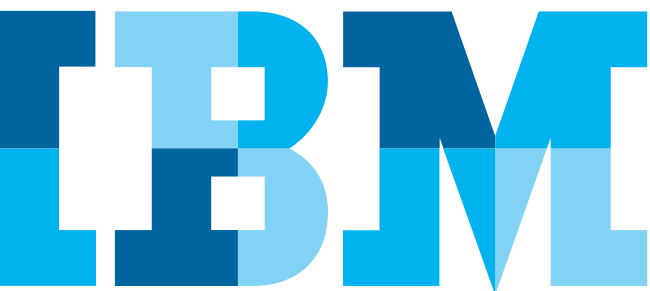
- Identify data in place that is relevant for legal collections or regulatory retention.
 - Act on data in an audited and defensible manner.
 - Implement and optimize use of tiered storage solutions.
 - Execute records retention and disposition policies.
 - Proactively implement a repeatable eDiscovery process.
 - React more quickly and with confidence to litigation requests.
 - Consistently adhere to data compliance and governance regulations.
-

Creating a data-driven enterprise: transforming unstructured data into relevant data

Each use case involving Big Data is a “volume-to-relevance” discovery process. That is why IBM® Unstructured Data Identification and Management is designed to help corporations discover, analyze, and act on terabytes to petabytes of unstructured data – providing a holistic solution for addressing the infinite data management problems that challenge records management, eDiscovery, compliance, storage optimization, and data migration initiatives. By providing an in-depth assessment of unstructured data across the enterprise, where it resides, IBM can provide organizations critical visibility into Big Data to make more informed business and legal decisions – turning unstructured data into relevant data.

Manage data in-place – the key to managing big data

The remarkable growth in unstructured enterprise data requires a new data management approach. Rather than move subsets of data into a repository or application database to manage it, IBM Unstructured Data Identification and Management moves the management to the data.



This is a mature approach using indexing technologies to understand and manage data in-place and provides speed, scale, reliability and performance of information management. IBM Unstructured Data Identification and Management provides:

- A highly efficient and intelligent indexing engine that dynamically gathers intelligence about data across the enterprise and stores this intelligence to be made available for virtually any business process.
- Massive scalability that can process immense amounts of data where it resides, while significantly reducing the impact on knowledge workers and corporate networks and servers.
- Distributed architecture, which links together or federates instances of IBM Unstructured Data Identification and Management to manage a company's data and optimize its location, routing, and delivery of the right information.
- Rapid deployment, in many cases, enables a company to quickly begin solving Big Data problems.

By replacing “moving and storing” as the governing principle with an “index and manage-in-place” strategy, enterprises can now manage their Big Data environments in much the same way they manage other IT networks ... simply and cost-effectively. IBM Unstructured Data Identification and Management enables an organization to discover, recognize, and act on unstructured data – without moving it to a repository or specialty application.

Discover: Big Data requires big search

IBM Unstructured Data Identification and Management accelerates the understanding of large amounts of unstructured content through powerful search functionality. The solution provides advanced search capabilities specifically tailored to assist legal, records, compliance, and IT staff who need to find responsive files in large amounts of unstructured data across the enterprise. No local application software is necessary to view searched file content, and viewing documents does not alter system metadata attributes.

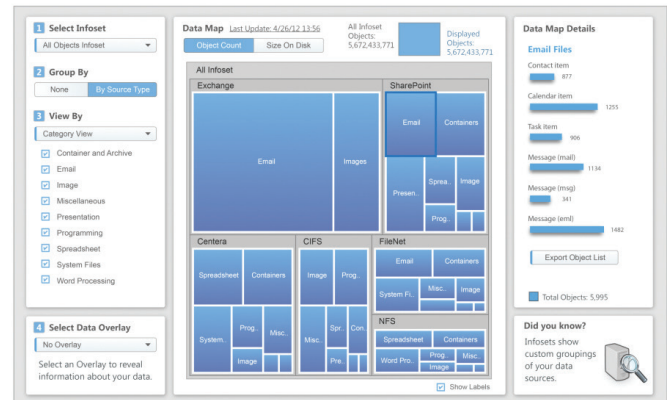


Figure 1: Advanced visualizations show what types of data are stored across the enterprise

The search functionality enables a user to enter simple to complex search strings and perform searches on an arbitrary mix of full text, metadata, and file attributes. IBM Unstructured Data Identification and Management supports the following advanced search capabilities:

- Keywords
- Single term search
- Multiple term search
- Boolean operators
- Wildcards in search terms or Boolean expressions
- Proximity searches
- Regular expressions
- Object level attributes
- Hash value searches (digital signatures)

Recognize: detailed data analysis increases corporate data intelligence

IBM Unstructured Data Identification and Management simplifies the analysis of unwieldy amounts of corporate data. IBM's ability to analyze and manage data in-place enables detailed analysis sooner and can limit impact on user productivity. This architecture enables IBM Unstructured Data Identification and Management to scale up to large environments without being limited to the storage capacities associated with repositories. The solution has the ability to present a variety of graphical perspectives of the information based on its metadata for simplified and efficient analysis and culling. Additionally, the entire contents of each data object can be displayed to gain thorough context of the information as necessary.

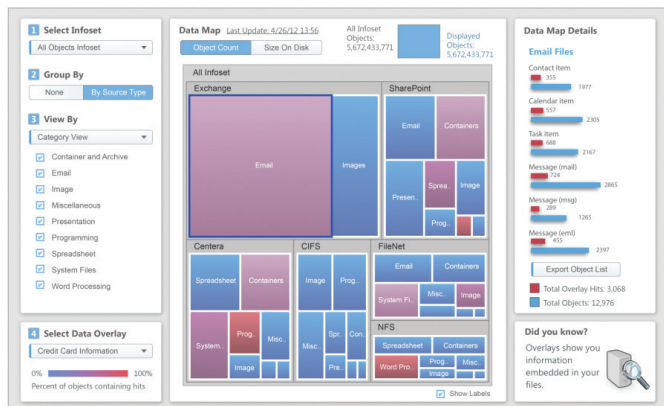


Figure 2: Utilize intelligent overlays to spot potential issues

IBM Unstructured Data Identification and Management provides a rich set of intuitive data explorers that provide visual representations of unstructured data from various points of view, such as file system metadata attributes (owner, size, date, type, etc.), data source location, and category, so stakeholders can better understand what data they have and what actions to perform. These explorers can be dynamically navigated to expose more details as necessary. IBM Unstructured Data Identification and Management also produces rich graphical reports that include the location of electronically-stored information (ESI) and the number of responsive and relevant documents.

Act: putting the power to act in your hands

What is the use of finding relevant data if you can not do something with it? IBM Unstructured Data Identification and Management provides support for many different policy actions such as copy, delete, move, copy to retention or export – each with a detailed audit trail. Whether your goal is optimizing storage, cleaning up your unstructured data, protecting corporate IP or customer information, retaining corporate records, legal collections, or virtually any number of other use cases, IBM Unstructured Data Identification and Management can provide you with the power to make informed decisions as you start down the path to true, active information management.

IBM Unstructured Data Identification and Management includes:

- **Data topology and overlay hit reports:** provide valuable insight and information that enables companies to make informed decisions prior to initiating an information governance initiative. The reports provide a snapshot of a client’s data environment based on location, server, users, etc.
- **Data filters:** enable clients to have coarse or fine grain control over their data to reduce data into smaller more meaningful or relevant information sets (i.e., “data over 7 years old”, “HR data located in the NYC office”, “custodians”, “matter terms”, etc.). These information sets can then be filtered further or acted upon, based on records retention policies, legal process, or IT protocol.
- **Actions:** enables the user to perform use case-specific functions such as “delete”, “move”, “copy”, or “export” data distributed over a variety of data sources, in a single unified manner. By using the IBM Unstructured Data Identification and Management distributed architecture, data that is separated across even large physical distances can be acted upon in a single motion to more quickly and easily facilitate a data governance project.
- **eDiscovery dashboard:** provides legal process, and workflows for early data assessment, collection, and export.
- **Incremental data collection:** files that are new or that have been modified since the last collection, are the ones that will be collected for preservation, helping ensure a streamlined and efficient collection process. This also helps ensure that the same document is not collected multiple times.
- **Auditability:** provides a detailed and thorough audit trail helping ensure object metadata and ownership is preserved with the time and date an action occurred and a forensic history that is designed to be sound and defensible. Audit information can be searched and exported for review by outside legal staff, if necessary. And, throughout the process, a virtually invulnerable chain of custody is maintained.

Summary

Unstructured data often represents significant legal cost and potential risk to an organization. Providing a powerful and highly scalable solution for data identification and management, IBM can help clients to gain visibility and control over enterprise data. IBM Unstructured Data Identification and Management is designed from the ground up to deliver on the manage-in-place philosophy. Built around a distributed architecture, IBM Unstructured Data Identification and Management can be quickly and easily deployed so that corporations can start with a small project and scale as their data grows, project scope increases, or new use cases emerge.

Why IBM?

Software and solutions are helping to build a smarter planet, giving organizations more opportunities to realize their potential and break new ground. To achieve such goals, businesses need solutions that are fueled by expertise, built for today's dynamic environment and designed to deliver tangible business results. IBM Information Lifecycle Governance Solutions can help. The robust ILG portfolio is designed to enable businesses to apply industry leading governance practices to both structured and unstructured content for retention, optimize business processes, maximize IT infrastructure and address complex compliance requirements. The deep ILG capabilities support an information economics agenda – the IBM approach to managing information as a strategic asset through better business alignment and coordination across internal stakeholder groups.

About IBM Information Lifecycle Governance Solutions

IBM's Information Lifecycle Governance (ILG) solutions help organizations improve their information economics by lowering information costs and risks while maximizing data value. IBM is one of the only providers of ILG solutions for Defensible Disposal, including holistically integrated solutions for efficient, rigorous eDiscovery, value-based retention and archiving, and governance and disposal enablement. These solutions are designed to enable legal, IT, records, and business users to reduce legal risk, data management and discovery costs. IBM ILG customers manage enterprise information by its business value; can comply more efficiently with litigation and regulatory duties, and dispose of information that has no value or duty.

For more information

To learn more, please contact your IBM representative or IBM Business Partner, or visit: ibm.com/ilg



© Copyright IBM Corporation 2014

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
January 2014

IBM, the IBM logo and ibm.com are trademarks of International Business Machines Corporation in the United States, other countries or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or TM), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. Other product, company or service names may be trademarks or service marks of others. A current list of IBM trademarks is available at “Copyright and trademark information” at: ibm.com/legal/copytrade.shtml

The content in this document (including currency OR pricing references which exclude applicable taxes) is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary. THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NONINFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



Please Recycle