



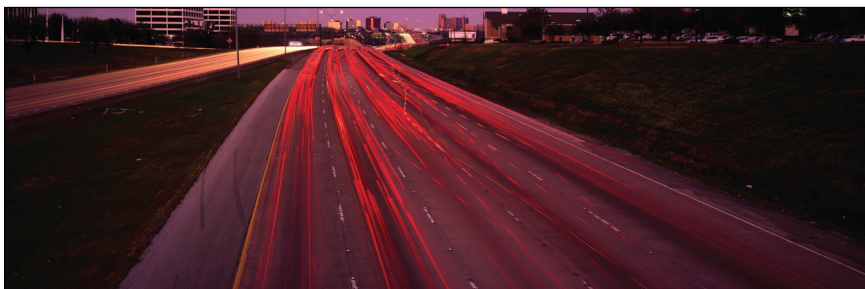
Information Management software

IBM InfoSphere Content Collector

Stop just archiving, start using your information

Highlights

- **Helps gain control over runaway storage and infrastructure costs**
- **Collects and archives content from e-mail, file systems, Microsoft SharePoint systems and more**
- **Employs a modular, extensible architecture that lets you add new content types and capabilities to meet new business requirements**
- **Provides capabilities for leveraging advanced classification, records management, eDiscovery and content-centric BPM capabilities as part of an ECM platform**
- **Uses smart archiving and content deduplication capabilities to help further reduce IT infrastructure costs**
- **Helps derive trusted information from unmanaged, unindexed, unnecessary “content in the wild” and bloated production systems**



Optimizing the value of your organization's content

Enterprise information continues to explode across three dimensions—volume, velocity and variety.

- **Volume**—Every day, 15 petabytes of new information are generated. By 2010, the codified information base of the world is expected to double every 11 hours.
- **Variety**—As much as 80 percent of new data growth is unstructured content, generated largely by e-mail, with increasing contribution by documents, images, video and audio.
- **Velocity**—The amount of information created has long surpassed storage capacity, and it continues to spiral beyond control. An average company with 1,000 employees spends US\$5.3 million a year to find information stored on its servers. And 42 percent of managers say they use the wrong information at least once per week.¹

In today's competitive business environment, you can't afford to drop your e-mail and other content into a siloed archive and never see it again. You need to use your content for a business advantage—by gaining control of information to better meet your legal obligations, connecting content so you can manage it together, integrating content with your business processes and lowering your total cost of ownership (TCO). You need to better manage your storage space. And you need to replace the common jumble of piece-part point products with modular, integrated solutions.

Delivering more capabilities than point e-mail archiving solutions

IBM InfoSphere™ Content Collector software is at the forefront of content collection and archiving solutions and collects, enhances and manages e-mail and multiple content types. With InfoSphere Content Collector, you can do more than just archive your information—you can start using it for business advantage.

More than e-mail: collect

The InfoSphere Content Collector approach enables you to monitor, identify and collect virtually all electronic content types from all locations including casually created content and information in existing archive silo systems. A significant advance over conventional e-mail archiving point solutions, InfoSphere Content Collector can collect—in addition to e-mail—Microsoft® file systems, collaborative documents created in the Microsoft SharePoint environment and it extends archiving capabilities to IBM Optim™, SAP and other software.

More than archiving: enhance

InfoSphere Content Collector does more than simply move your content. It accesses, monitors and enhances content by tasking, filtering, transforming, extracting metadata, customizing and more—including multiple classification options—with no costly custom integration required and with no burden on the end user. InfoSphere Content Collector is powered by IBM ZeroClick technology, which uses process, events, rules, metadata and advanced classification to transparently enforce policy throughout the entire compliance information life cycle.

More than an isolated solution: manage

With InfoSphere Content Collector, you can manage content as part of an enterprise content management (ECM) platform and information infrastructure—not a silo—to unlock value while leveraging existing investments. Importantly, its extensible architecture leverages investments and best practices from ECM systems from IBM, without requiring complex services engagements.

Enhancing flexibility with a modular, extensible architecture

The ability of InfoSphere Content Collector to handle a range of content types and sources and to add new capabilities as needed is based on the solution's modular, extensible architecture. A flexible set of source connectors, task connectors and target connectors can be mixed and matched to create the functionality that users require from the technologies they already have installed.

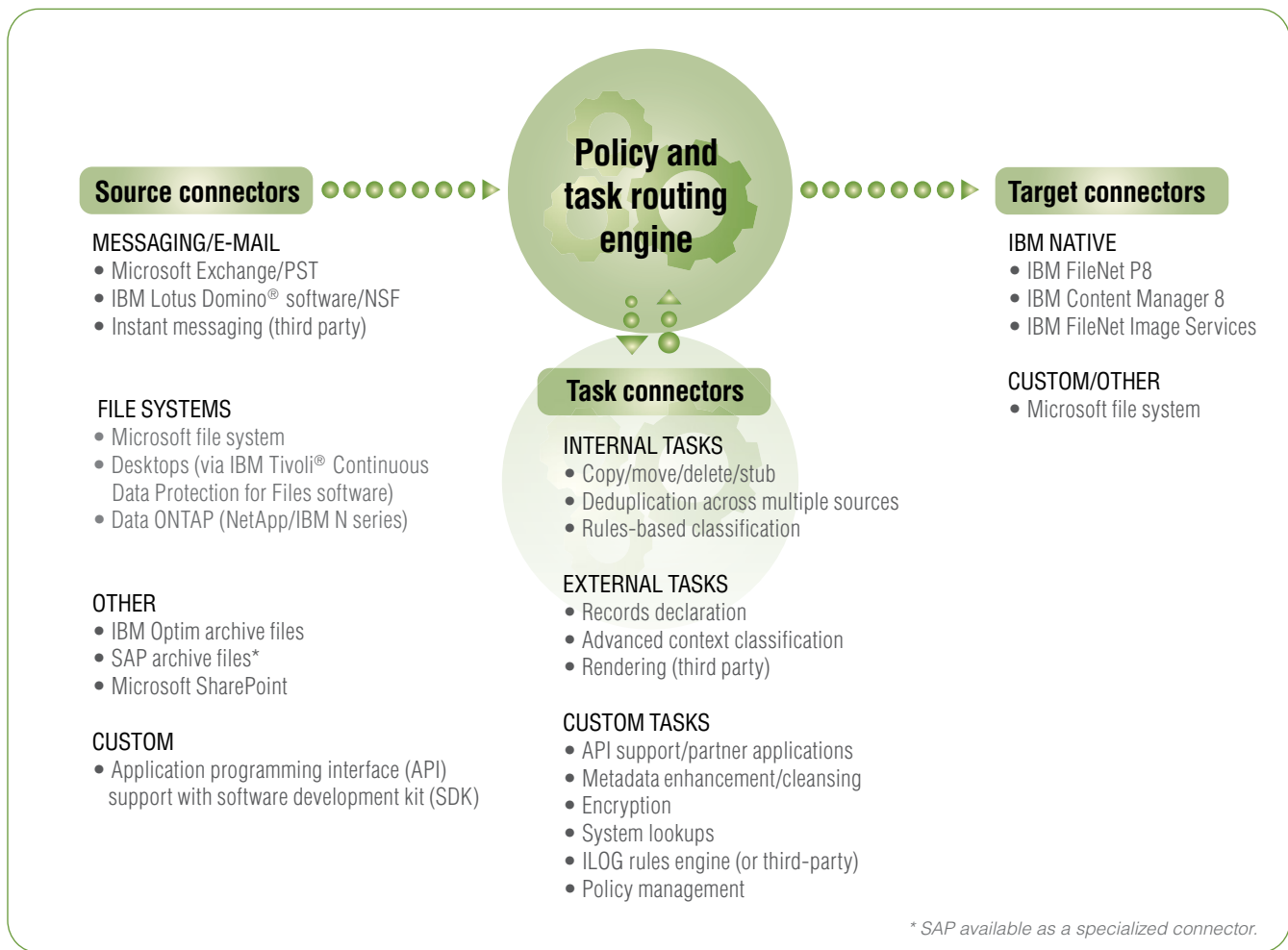


Figure 1: Content collection and archiving offerings from IBM feature a modular, extensible architecture that provides the flexibility to grow with your needs.

Integrating with a comprehensive portfolio of compliance tools

InfoSphere Content Collector integrates with IBM InfoSphere Classification Module, IBM InfoSphere Enterprise Records, IBM InfoSphere eDiscovery Manager and IBM InfoSphere eDiscovery Analyzer software. It helps you implement a comprehensive solution that includes advanced classification, records management, electronic discovery (eDiscovery) searches and analytics for the entire compliance information management life cycle—and it provides tools you can use to manage compliance-related information in a security-rich, auditable manner.

Meeting content collection and archiving needs with comprehensive capabilities

IBM InfoSphere Content Collector offers multiple solutions to best suit your content collection and archiving needs:

- *IBM InfoSphere Content Collector for Email software, for e-mail collection and archiving*
- *IBM InfoSphere Content Collector for File Systems software, for file system collection and archiving*

- *IBM InfoSphere Content Collector for Microsoft SharePoint software, for Microsoft SharePoint document collection and archiving*
- *IBM InfoSphere Content Collector Discovery Analytics software, an all-in-one offering for content collection and archiving of e-mail and file-system content plus eDiscovery and analytics*

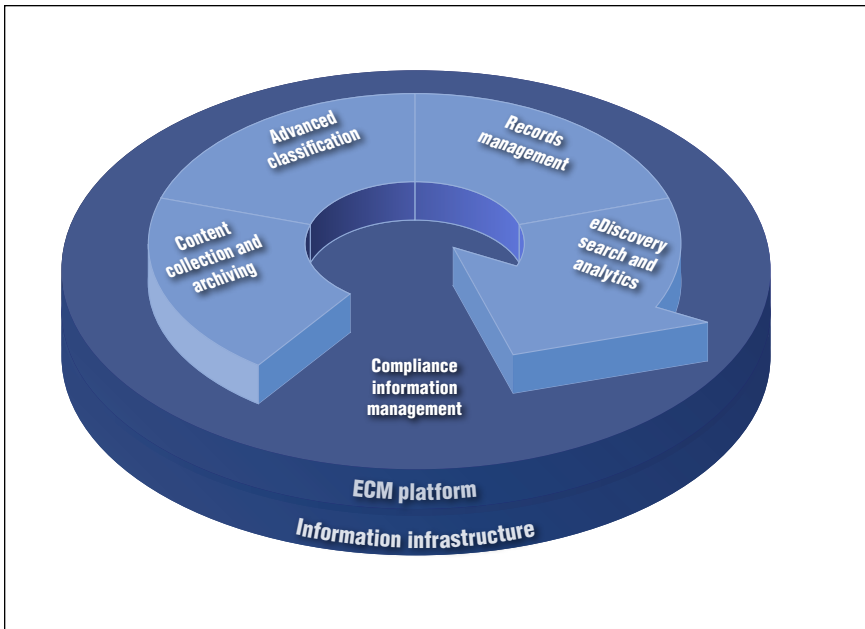


Figure 2: The content collection and archiving capabilities of InfoSphere Content Collector software integrate with advanced classification, records management and eDiscovery search and analytics capabilities in your ECM environment. They also integrate with IBM information infrastructure solutions designed to deliver improved service levels, reduce business risks and manage today's rapidly growing volumes of business information. The result—optimized processes throughout the life cycle of compliance information management.

Common capabilities

All InfoSphere Content Collector offerings share powerful capabilities that can help you:

- *Simplify and enhance configuration.*
 - Shortens initial installation and base configuration to as little as one hour, not including source systems or target repository
 - Enables task routes to be easily configured, reviewed and modified via a configuration designer
 - Allows extending the modular architecture by adding custom sources and tasks
- *Reduce organizational risk, ensure consistency, lower cost and improve productivity with powerful enhancement options enabled by ZeroClick technology, including:*
 - Automated, policy-driven collection and archiving of e-mail and content
 - The ability to base policies on metadata such as sender, sent date, e-mail domain, or file folder structure or location
 - Keyword matching and pattern matching
 - The ability to plug into multiple forms of classification including IBM InfoSphere Classification Module
 - The ability to declare content a record with IBM InfoSphere Enterprise Records software

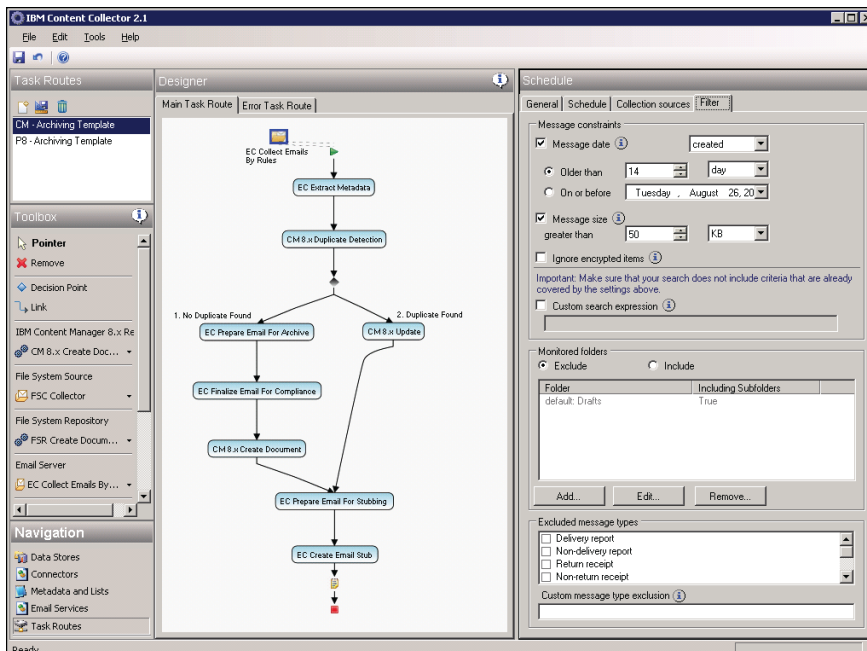


Figure 3: InfoSphere Content Collector software task routing supports more than archiving by enabling flexible routing and configuration for processing of internal and external instructions, or tasks.

- Options for manual collection and archiving for e-mail and other content
 - The ability to capture metadata and use it with other ECM capabilities including eDiscovery, records management and business process management (BPM)
 - Options for leaving source e-mail and content in place, moving it, flagging it as captured, stubbing it or deleting it
 - Deduplication of e-mail attachments and other documents collected from various source connectors
 - Content retention and disposition management to apply consistent retention policies to your content
- Manage e-mail for legal obligations.
 - Provides automated, policy-driven collection and archiving using ZeroClick technology, with journaling and manual options also available
 - Provides capabilities for dynamic records declaration with InfoSphere Enterprise Records
 - Enables eDiscovery of e-mail using InfoSphere eDiscovery Manager and InfoSphere eDiscovery Analyzer tools
 - Connect e-mail to managed content.
 - Manages e-mail along with other managed content in either IBM Content Manager or IBM FileNet P8 software
 - Enables users to optionally add additional metadata during manual e-mail capture operations
 - Manage e-mail as part of a business process.
 - Automatically captures e-mail from monitored mailboxes and inserts them into business processes in FileNet P8 software; also allows the option of manual capture
 - Archive e-mail for storage space management.
 - Reduces storage space on e-mail servers with mailbox life-cycle management tools including stubbing and deduplication
 - Enables owners to manage their mailbox size by flagging e-mails to be archived
 - Reduce the management burden and improve productivity of end users.
 - Supports IBM Lotus Notes® and Microsoft Exchange e-mail software
 - Allows mailbox owners to select e-mails for archiving
 - Enables users to search for e-mails in the repository based on sender and recipient information, date ranges and content
 - Enables users to restore e-mails to their in-box or to restub them
 - Enables users to view archived e-mail in their environment of choice—native IBM Lotus Notes, IBM Lotus® iNotes®, Microsoft Office Outlook, Microsoft Outlook Web Access software or in a Web browser
 - Provides users with offline access to their archived e-mails even when not connected to the repository
 - Supports collection and archiving of Microsoft Outlook PST files and public folders
 - Supports collection and archiving of Lotus Notes databases

IBM InfoSphere Content Collector for Email, Version 2.1.1 software

An automated, flexible and scalable answer to the challenge of managing e-mail, this offering supports IBM Content Manager, Version 8 and IBM FileNet® P8 software. It provides tools designed to help you:

- Simplify and enhance configuration.
 - Allows configuring multiple servers in a single instance using a single administration console, simplifying administration and enhancing high-availability and failover support

IBM InfoSphere Content Collector for File Systems, Version 2.1.1 software

An enterprise solution supporting the Microsoft Windows NT® file system (NTFS) format for controlling documents and content on network file-share drives, this offering supports Content Manager, Version 8, FileNet P8 and IBM FileNet Image Services software. It also supports document handling procedures that comply with corporate and regulatory policies by providing tools that help you to:

- *Manage files for legal obligations.*
 - *Classifies documents and content using metadata, rules or IBM Classification Module software*
 - *Allows changing file system security to read only at time of capture to help maintain document authenticity*
- *Connect disconnected files to other managed content.*
 - *Automatically captures documents as they are placed in a monitored location on the file share*
 - *Enables leaving documents in place, or moving, flagging or replacing them with shortcuts so users can find previously loaded documents*

- *Manage files as part of a business process.*
 - *Connects content not in the ECM repository to business processes throughout the enterprise*
 - *Permits sophisticated and targeted file handling using an advanced rules engine*
 - *Allows metadata in comma-separated value (CSV) and XML files to be automatically captured in the repository and operated upon*

IBM InfoSphere Content Collector for Microsoft SharePoint, Version 2.1.1 software

This offering can help address content management requirements essential for IT and compliance policy enforcement, and at the same time provide an integrated ECM solution that incorporates Microsoft SharePoint software. InfoSphere Content Collector for Microsoft SharePoint software can help you leverage broader BPM services and long-term content storage and retention from a master content management standpoint.

InfoSphere Content Collector for Microsoft SharePoint software supports sweeping content from Microsoft Windows® SharePoint Services (WSS) 3.0 or Microsoft Office SharePoint Server 2007 document libraries into Content Manager, Version 8 or FileNet P8 repositories, based on administrator-configured business rules. This content can be copied (optionally marking the documents as read-only or deleted from SharePoint) or moved to the IBM FileNet P8 repository (optionally leaving a link to the FileNet P8 object in the Microsoft SharePoint document library).

As content is swept from Microsoft SharePoint software, tasks can be assigned that can perform a wide variety of additional operations, including:

- *Adding metadata to the content.*
- *Determining which item type/object store and security features to use.*
- *Deduplicating redundant content.*
- *Setting file plans or declaring documents as records.*
- *Allowing customers or IBM Business Partners to build their own custom tasks.*

IBM InfoSphere Content Collector

Discovery Analytics, Version 2.1.1 software

Organizations can now find all the benefits and capabilities of InfoSphere Content Collector for Email and InfoSphere Content Collector for File Systems, plus reap the advantages and cost savings of eDiscovery with InfoSphere eDiscovery Manager and InfoSphere eDiscovery Analyzer software all in one offering. InfoSphere Content Collector Discovery Analytics software can:

- *Jump-start your small to midsize projects with 250 to 1,400 users with the option to purchase a starter pack.*
- *Provide eDiscovery capabilities that can help your organization respond to discovery demands quickly and cost-effectively.*
- *Collect and preserve electronically stored information in a security-rich, traceable, high-scale repository that includes chain-of-custody tracking.*
- *Help your organization strategically control the explosive growth of information, helping reduce the costs and risks of eDiscovery.*

Offering solutions for the ECM

installed base

If you already have an ECM solution from IBM, the modular, extensible architecture of InfoSphere Content Collector software supports existing investments in IBM systems—either IBM Content Manager, Version 8; IBM FileNet P8, Version 4; or IBM FileNet Image Services software in the case of InfoSphere Content Collector for File Systems software. The architecture enables organizations using IBM CommonStore² or IBM FileNet Email Manager software to easily add Microsoft SharePoint and file system management capabilities.

And it enables organizations using IBM FileNet Records Crawler software to easily add e-mail and Microsoft SharePoint content management capabilities.

Meeting the needs of non-IBM users

If you don't already have an ECM solution from IBM, InfoSphere Content Collector software together with Content Manager, Version 8 or FileNet P8 software can help you gain control and unleash the value of your content. The modular and extensible source and task connectors in InfoSphere Content Collector software also have the flexibility to integrate with other existing solutions.

Why IBM?

IBM InfoSphere Content Collector software is a key part of the IBM strategy for a smarter approach to archiving, which provides capabilities organizations can use to:

- *Gain a deeper understanding of what information to archive through discovery- and analytics-based assessment technologies.*
- *Eliminate point solution complexity and cost by unifying data and content archiving through common collection (ingest) technologies that create an information-aware infrastructure.*
- *Implement cost-optimized retention leveraging a unified, flexible and security-rich information infrastructure.*
- *Speed time to value through multiple modular yet integrated solutions including choice of management and delivery models based on a common information lifespan and policies.*
- *Reduce risk, respond faster to legal inquiries, and trust and leverage information through value-added capabilities.*
- *Optimize and protect information with system and managed service offerings.*



ECM solutions from IBM enable the world's top companies to make better decisions, faster. They deliver a broad set of mission-critical solutions that help solve today's difficult business challenges: managing unstructured content, optimizing business processes and helping to satisfy complex compliance requirements. More than 13,000 organizations rely on ECM solutions from IBM to improve performance and remain competitive through innovation.

Ordering information

IBM InfoSphere Content Collector for Email per user value unit LIC+SW S&S 12 MO: part number D05XQLL

IBM InfoSphere Content Collector for File Systems per user value unit LIC+SW S&S 12 MO: part number D05XNLL

IBM InfoSphere Content Collector for Microsoft SharePoint per user value unit LIC+SW S&S 12 MO: part number D0AYNLL

IBM InfoSphere Content Collector Discovery Analytics per user value unit LIC+SW S&S 12 MO: part number D0ARGLL

For more information

To learn more about IBM InfoSphere Content Collector software and other ECM solutions from IBM, contact your IBM representative or IBM Business Partner, or visit:

ibm.com/software/ecm

For additional details about IBM InfoSphere Content Collector software—including hardware and software requirements—visit:

ibm.com/software/data/content-management/content-collector

© Copyright IBM Corporation 2009

IBM Corporation
Software Group
3565 Harbor Boulevard
Costa Mesa, CA 92626-1420
U.S.A.

Produced in the United States of America
October 2009
All Rights Reserved

IBM, the IBM logo, and ibm.com are trademarks or registered 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 ™), 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. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

Microsoft, Windows, and Windows NT are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. The information contained in this documentation is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this documentation, it is provided "as is" without warranty of any kind, express or implied. In addition, this information is based on IBM's current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this documentation or any other documentation. Nothing contained in this documentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM (or its suppliers or licensors), or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

Each IBM customer is responsible for ensuring its own compliance with legal requirements. It is the customer's sole responsibility to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

1 Steve Delahunty, "State Of Enterprise Storage: Changing Priorities, Changing Practices," *InformationWeek Analytics*, January 2009.

2 Certain specialized products, including CommonStore for SAP, are not changing architectures and will retain the CommonStore name.