

# EMC Documentum Content Intelligence Services

## Automatic content classification helps you locate and reuse information

### THE BIG PICTURE

- Classifies large volumes of content automatically
- Enables a combination of automated and manual classification
- Includes pre-built, industry-specific taxonomies to ease the process of taxonomy creation
- Supports classification of content in its native language

There's little argument over the value of enterprise information. Organizations that leverage their information—or content—get products to market faster, service their customers better, and boost operating margins by streamlining inefficient business processes.

But today, many organizations are challenged by the enormous volume and rapid growth of enterprise content. That content comes from inside and outside the organization in many forms—documents, XML components, HTML pages, reports, and scanned images. This information is crucial to business performance and yet, to leverage it effectively, it must be well organized. Like a library with millions of books and no organizational structure, electronic content is virtually useless without a systematic method for indexing, cataloging, and retrieving it. And this is where EMC® Documentum® Content Intelligence Services (CIS) can help.

### The power of CIS

Content Intelligence Services is an extension to the EMC Documentum content management platform that enables automatic classification and categorization of content in the Documentum repository. Its benefit is content that is well organized, properly classified, and correctly categorized—facilitating reuse and allowing users to navigate content through an intuitive structure.

### Automated classification

The power of automated classification lies in processing large volumes of information quickly—as it already exists in the repository or as it arrives from authoring applications, record filing, or business processes. With CIS, content is parsed and analyzed and classification rules are applied. The results of the classification can then be used for categorization or as keywords to populate content metadata.

The capability of automatically creating keywords to populate content metadata can remove the burden from end users who otherwise must remember to do it manually. Many users struggle to consistently populate metadata as content is being created, which significantly limits its future use since metadata is what enables programmatic processing of content. CIS eliminates this dependency on end users. EMC Documentum Web Publisher, for instance, uses CIS to automatically propose metadata (metatags) to the web master who can simply accept them or modify them as needed.

## Prepackaged taxonomies

To ease the process of taxonomy creation, CIS includes prebuilt taxonomies for the following industries and corporate functions:

- Automotive
- Chemicals
- CRM
- DOE
- Drug/Disease
- Energy
- Engineering
- European Union Vocabulary
- Federal Legislation
- Financial
- Genes
- Geography
- Global Law
- Healthcare
- High Tech
- Human Resources
- Human Rights
- I T
- Legal
- Marketing
- Medical / Health (CRISP)
- MEDDRA
- Military
- Military
- Petroleum
- Pharmaceutical
- Physics

## Combining automated and manual classification

Many organizations are not comfortable with a complete reliance on automatic classification. They want to harvest the competence of various teams and individuals while fine tuning the automatic classification process. CIS provides support for a combination of automatic and manual classification with a special user interface for category owners. Category owners can make classification decisions manually in cases where the automated rules cannot classify content with a preset certainty level. The user interface for category owners is built into every Documentum client such as Webtop and becomes active upon detection of CIS in the system.

## Taxonomy-based categorization

With CIS, the results of classification can be used for content categorization, which assigns content to appropriate categories. Typically, categories are represented by portal navigation or by a folder structure to which content is linked. A category hierarchy—or taxonomy—is usually common to an entire department or organization and allows for all users to share the same navigational view for content in active projects or content that has been archived or filed as records.

The creation of categories from scratch can be a massive task requiring many functions in the organization to collaborate. To ease this process, CIS comes with prepackaged taxonomies for various industries and roles. These taxonomies can be customized and used either out of the box or as a starting point for a customization. Category owners can also leverage their expertise and add categories and sub-categories in their areas.

## Multilingual support

Content Intelligence Services natively supports major European languages including English, German, French, Spanish, Portuguese, Dutch, Swedish, Norwegian, Danish, and Finnish. This enables the classification of local content in its native language against an enterprise-wide or local taxonomy. Using these multilingual capabilities, companies can deploy CIS globally, enhancing the globalization capabilities of Documentum that include pervasive Unicode compatibility and localized user interfaces.

## About EMC

EMC Corporation (NYSE: EMC) is the world's leading developer and provider of information infrastructure technology and solutions that enable organizations of all sizes to transform the way they compete and create value from their information. Information about EMC's products and services can be found at [www.EMC.com](http://www.EMC.com).

# EMC<sup>2</sup>

where information lives®

EMC Corporation  
Hopkinton  
Massachusetts  
01748-9103  
1-508-435-1000  
In North America 1-866-464-7381  
[www.EMC.com](http://www.EMC.com)

### Take the next step

To learn more about EMC Documentum Content Intelligence Services, visit [www.EMC.com](http://www.EMC.com) or call 800.607.9546 (outside the U.S.: +1.925.600.5802).