

# Automatic Content Categorization and Tagging with Content Intelligence Services

*July 2003*

## Contents

Executive Summary.....	3
Business Challenges — and the Benefits of Intelligent Content .....	4
Inability to Find the Right Information .....	4
Disorganized Information.....	4
Inconsistent, Manual Tagging.....	4
Inability to Personalize Content.....	4
Consequences.....	5
What's the Solution?.....	5
CIS: Structuring Unstructured Enterprise Content .....	6
Balancing Automation and Manual Processes .....	7
Overview of CIS Capabilities .....	7
Terminology and Key Concepts.....	8
CIS Architecture .....	9
Information Extraction.....	10
Conceptual Classification .....	11
Semantic Analysis .....	12
Tagging and Categorization .....	13
Pre-built Starter Taxonomies.....	15
Taxonomy Exchange Format (TEF).....	15
Application Examples .....	16
Personalized Web Sites and Cross-Selling .....	16
Pharmaceutical Clinical Trial Sites .....	16
Business News and Event Tracking .....	17
Content/Component Reuse Support .....	17
Business Benefits .....	17

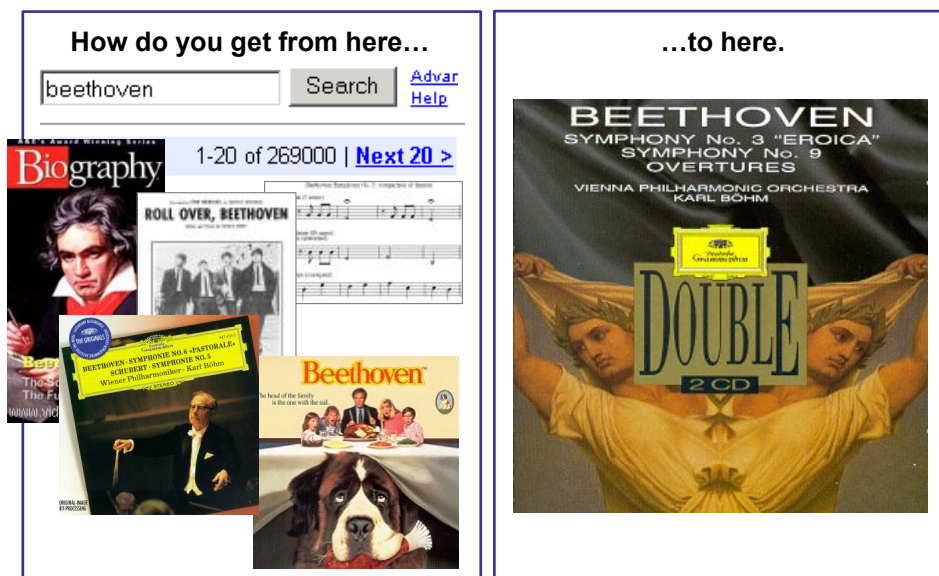
## Executive Summary

Industry analysts suggest that over 80% of the content within the typical Global 2000 organization is unstructured – that is, content which does not fit neatly into the rows and columns of a relational database. The fact is, this unstructured content is key to many business processes across the organization and throughout the business value chain, from engineering drawings and specifications, to brand assets such as sales and marketing collateral, legal documents, educational videos and tutorials, online product catalogs, and customer service information. With each new product or service your company delivers to the market, with each new initiative you launch, and with each new Web page that is created, this unstructured content is growing at an exponential rate.

At the same time, your knowledge workers need quick, seamless access to this information, which can often be a daunting task. In fact, employees spend over 40% of their time just looking for the information they need to effectively do their jobs, and are often frustrated because the information they need cannot be accurately found, the latest version is not available, or they don't even know that it exists (Forrester Research).

As a result, enterprise content gets recreated instead of being repurposed and reused. Documentum Content Intelligence Services (CIS) addresses the inability to organize and repurpose unstructured enterprise content by automating what are otherwise manual, often inconsistent content tagging and categorization processes and applying advanced metadata generation capabilities. In short, Content Intelligence Services organizes business-critical content by automatically tagging and categorizing it — turning unstructured content into intelligent, structured content.

With these powerful capabilities, CIS enables precise searching, easy navigation, and effective personalization while promoting content reuse across multiple initiatives and increasing productivity through process automation. And Documentum is the only vendor to offer these powerful content categorization and metadata capabilities fully integrated with an industry-leading content management platform. With Documentum CIS, content becomes more intelligent, allowing you to leverage business-critical information more effectively across multiple enterprise content management initiatives, including corporate Web sites and partner and employee portals. The bottom line: your content becomes perfectly clear — well organized, easy to find, tailored to individuals.



Making content more intelligent allows you to turn chaos into clarity.

## Business Challenges — and the Benefits of Intelligent Content

The explosion in information, or content, has created several business challenges, including:

- Inability of users to find the information they're looking for
- Lack of clear organization to simplify navigation through repositories and on Web sites
- Manual tagging processes that take far too much time to complete and are inconsistent at best
- Inability to easily personalize content for individual users and customers

### Inability to Find the Right Information

Your knowledge workers need on-demand, 24x7 access to enterprise information assets, but providing this access can often be a daunting task. In fact, employees spend over 40% of their time just looking for the information they need to effectively do their jobs. For example, think about how often you've done a keyword or full-text search for something on a Web site, corporate portal or repository that returned thousands of irrelevant items. Compare that to the number of times you *have* been able to find exactly what you are looking for.

As a result of inaccurate search capabilities, enterprise content gets recreated instead of being repurposed and re-used, and billions of dollars are spent in lost opportunities. Search engines are key to solving this problem, but without rich metadata such as keyword associations, the results are often irrelevant. You need to tag your content — and fast!

### Disorganized Information

In addition to the challenges associated with enabling accurate search of your enterprise content assets, you must be able to organize it all for easy navigation. But to accomplish this goal, you first need to create a logical folder structure that reflects the categories of content existing within your organization (sales, marketing, engineering, HR, IT), that includes your particular vertical market characteristics (life sciences, energy, high tech, financial services), and that encompasses business-specific relationships (competitors, partners, investors, customers). Although on a simplistic level, this folder structure represents categories and sub-categories that define your company's content, this "taxonomy" is often difficult and time-consuming to create. Developing an effective corporate taxonomy requires skilled resources such as taxonomy scientists and domain experts and can take several weeks or even months to develop.

### Inconsistent, Manual Tagging

In addition to creating an effective taxonomy structure, you must also generate and apply rich metadata or attributes to new content as it is created. This process is expensive to do manually. Although some IT organizations and Web teams create custom attributes for Web content, many leave the actual attributing to the individual content author or contributor on a one-off basis as content is saved into the corporate repository. This attribution process can vary across organizations and departments, often undoing much of the tagging benefit as attributes are erroneously or inconsistently applied to new or existing content.

### Inability to Personalize Content

Finally, you need to respond in a personalized manner to online customer, partner, and employee requests, and this means delivering the right content on demand, in the right context, in the right format, and through the right delivery channel. You can employ a personalization server, but to effectively match customer requests with profile information personalization servers require both content and rich metadata. In short, you need an automated process for generating and applying metadata, allowing you to engage in high value, personalized interactions with customers via the Web to boost revenues and customer retention.

## Consequences

Ultimately, the cost of not effectively addressing these challenges can lead to billions of dollars in lost business opportunity, including:

- Poor productivity as employees and partners spend too much time looking for business-critical information
- Excess costs as business-critical content gets re-created instead of reused
- Lost content that gets inconsistently tagged and categorized on an ad-hoc basis
- Unhappy, frustrated customers that are hesitant to do business with you the site's search engine is not returning relevant results, the navigation is not intuitive, and the content is not personalized in the proper context
- Stale Web content that is lacking the rich metadata needed for dynamic publishing
- Increasing deployment costs as excess time and resources are spent creating customized taxonomies from scratch

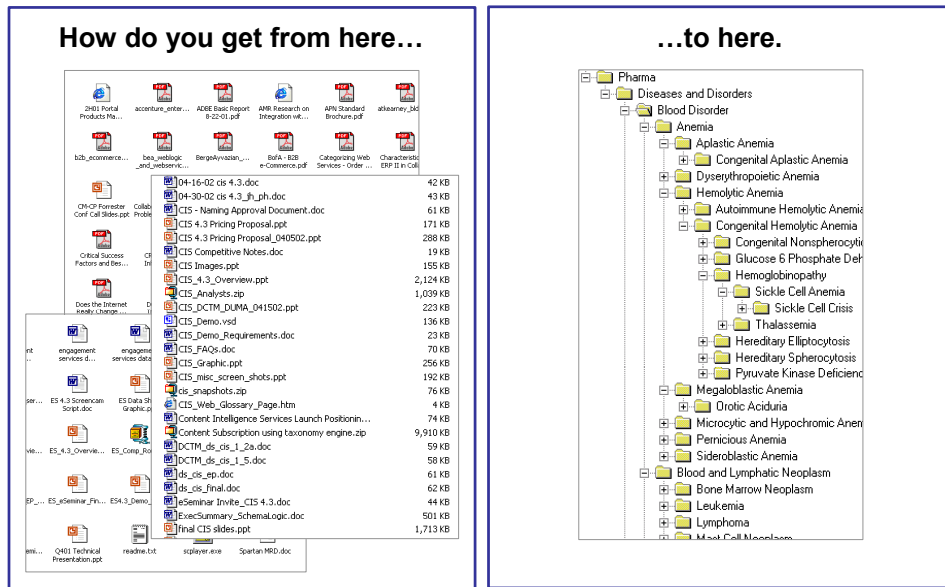
## What's the Solution?

To overcome these business challenges, you need an enterprise content management (ECM) system capable of managing all content types, including documents, XML, Web content, and rich media, and a content intelligence solution capable of:

- Automating the content tagging and categorization process with the right balance between automation and manual approval and management of metadata
- Enabling rich metadata extraction capabilities
- Conceptually classifying content to determine it's "aboutness"
- Providing pre-built, starter taxonomies based on industry-standard thesauruses to speed deployment
- Analyzing content according to business semantics to increase associations between content and drive additional business processes
- Seamlessly integrating with an ECM server and repository to support content reuse across multiple initiatives, leverage existing investments, and achieve lower total cost of ownership (TCO)

Documentum Content Intelligent Services provides these capabilities, enabling:

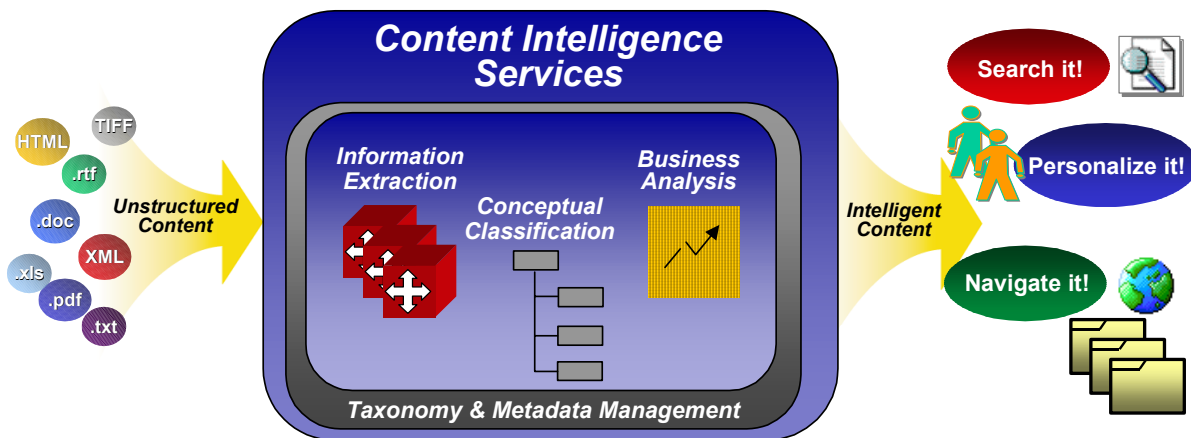
- Accurate searches
- Clear organization and simplified navigation
- Consistent, automated, and rapid content tagging
- Personalized content delivery tailored for individuals



Content Intelligence Services provides pre-built taxonomies and metadata management capabilities to organize content for easy navigation.

## CIS: Structuring Unstructured Enterprise Content

Documentum Content Intelligence Services (CIS) is an extension to the core Documentum Content Server that automates and controls the enrichment and organization of enterprise content based on powerful information extraction, conceptual classification, business analysis, and metadata and taxonomy management capabilities. In simpler terms, CIS takes unstructured content and adds attributes (or metadata) to turn it into intelligent, structured content to enable more accurate searching, easier navigation, and more effective personalization of enterprise content.



Documentum Content Intelligence Services adds structure to unstructured content to enable more accurate searching, easy navigation, and effective personalization.

## Balancing Automation and Manual Processes

CIS provides the right balance between automated and manual processes. Leveraging Documentum workflow and lifecycle management, you can add content intelligence as an automated step specifically for tagging and categorizing content, ensuring that this process becomes a standard best practice across the enterprise. For increased efficiency, you can run CIS during batch processing for tagging large volumes of content currently stored in the corporate repository. And as new content is added, it can be tagged and categorized automatically, based on pre-determined taxonomies.

While automation is generally preferred, there are times when you will want to return greater control to users through manual processes. With CIS and Documentum Web Publisher, you can set up attributes that are selected by content contributors when they create or add new content to a repository. Web Publisher can also be configured to suggest metadata values when content is checked in, allowing each user to select from the list of suggested or predetermined attributes. While manually tagging and categorizing content may lead to inconsistencies and errors with metadata that may not be in alignment with corporate standards and nomenclature, there are times nonetheless when these decisions are best left up to individuals.

## Overview of CIS Capabilities

More than just a simple program for adding metadata to content, Documentum CIS provides a wealth of highly valuable content intelligence capabilities. At a high level, CIS supports the following four areas:

- Information extraction
- Conceptual classification
- Business (semantic) analysis
- Taxonomy and metadata management

The capabilities supporting these four key areas are described in the following table.

CIS Capability	Business Benefit
Auto-tagging and auto-categorization	<p>Reduces costs by automating manual, inconsistent tagging and categorization efforts</p> <p>Maximizes content reuse and consumption and enables effective management of brand resources</p>
Business (semantic) analysis	Automatically analyzes content to reflect your company's business objectives and drive further business processes
Pre-built departmental and vertical taxonomy libraries	Reduces costs and speeds time to implementation by providing a base taxonomy model for a particular department or industry application
Standards-based XML taxonomy import utility	<p>Speeds implementation time by automatically defining the appropriate repository structure based on imported taxonomy</p> <p>Protects your investment in existing taxonomies through the implementation of industry standards (XML)</p>
Taxonomy and metadata management	Makes it easy for you to create, customize, and maintain taxonomies

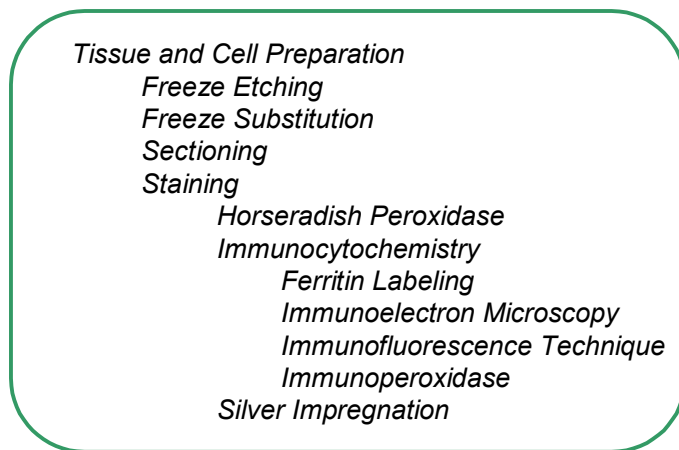
## Terminology and Key Concepts

Before we can discuss the technical details of how CIS delivers these capabilities, we first need to define some key terms: taxonomy and domain map.

A **taxonomy** is simply a listing of categories and sub-categories of content that may exist in your organization — a hierarchical structure of concepts. To effectively categorize your enterprise content, your organization may require several taxonomies: an industry-specific taxonomy that incorporates standard terminology (for example, an aerospace taxonomy that incorporates ATA terminology); departmental taxonomies (IT, HR, sales, marketing, finance); and a company taxonomy (with categories for products, competitors, investors, partners, and so on). Taxonomy concepts or categories are typically represented in a structured, hierarchical form where the nodes in this structure are referred to as “taxonomy nodes.” The hierarchical structure represented by the taxonomy nodes can be utilized for logical operations, including:

- Categorization of unstructured documents,
- Building a Documentum repository folder structure, and
- Aiding in categorical navigation

Within a taxonomy, all concepts are equal and the taxonomy node hierarchy is simply used for convenience. An example of a hierarchical taxonomy for life sciences is as follows:



Example life sciences taxonomy - several nodes deep.

**Domain maps** represent these same concepts along with the keyword evidence for each concept. Therefore, a domain map is simply a taxonomy plus the evidence associated with each of the concepts. It is common for the structure of the concepts represented in the domain map to match the structure modeled in the taxonomy to simplify administration. In other words:

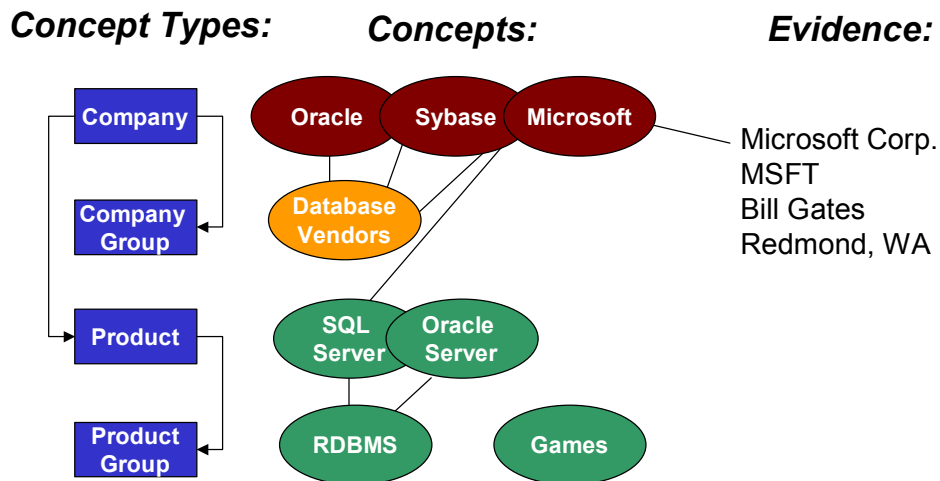
**Domain Map = Taxonomy + Evidence Terms**



CIS uses a domain map to analyze a document to locate keywords, and then uses a mapping of keywords to concepts to decide which is the most likely key concept for a given document. CIS administrators and developers can retrieve all concepts and their weighting based on the evidence found in the document. The three components included in the CIS domain map are *concept types*, *concepts*, and *evidence*.

The following domain map example shows four main categories of information: Company, Company Group, Product, Product Group. Each of these categories of information is considered a *concept type* in the domain map. Within each concept type, there are subcategories of information that describe each category; called *concepts*. The *evidence* proves the existence of each concept in a document.

## Domain Map Components



The CIS domain map incorporates business rules that link taxonomy nodes with evidence to enable calculated content auto-tagging.

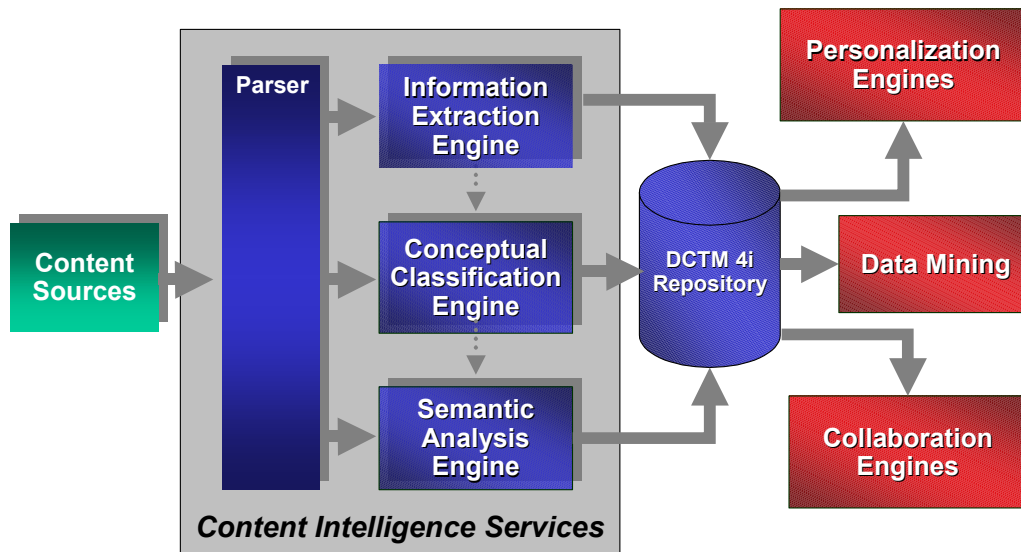
The existence of any of the following words or phrases infers that the document is about Microsoft (the *concept*): Microsoft Corp., MSFT, Bill Gates, and Redmond, WA. One advantage with the use of concept definitions is that CIS will not need the existence of the word Microsoft or SQL Server in the document in order to infer that the document is about Microsoft.

## CIS Architecture

CIS operates as a set of extended services for the core Documentum Content Server and repository, and consists of four key components: a **parser**, an **information extraction engine**, a **conceptual classification engine**, and a **semantic analysis engine**. Content sources to be tagged and classified are first parsed into smaller components (for example, documents parsed into words) by the CIS parser. After parsing, metadata is then extracted from the content for classification based on a taxonomy, pre-defined concepts, and powerful business rules. It is important to note that CIS processing is flexible in that not all of the subsequent processing steps are required after parsing (as indicated in the diagram below); however, for maximum benefit all processing steps should be utilized. Once CIS is configured, it can be scheduled to run automatically (based on a predetermined schedule, a workflow task, or

lifecycle state) to obtain documents from folders in the Documentum repository, create conceptual models, set document attributes, and link the documents to specific repository folders.

## Content Intelligence Services Architecture



Content Intelligence Services extends the core ECM services provided by the Documentum Content Server to automatically organize content.

### Information Extraction

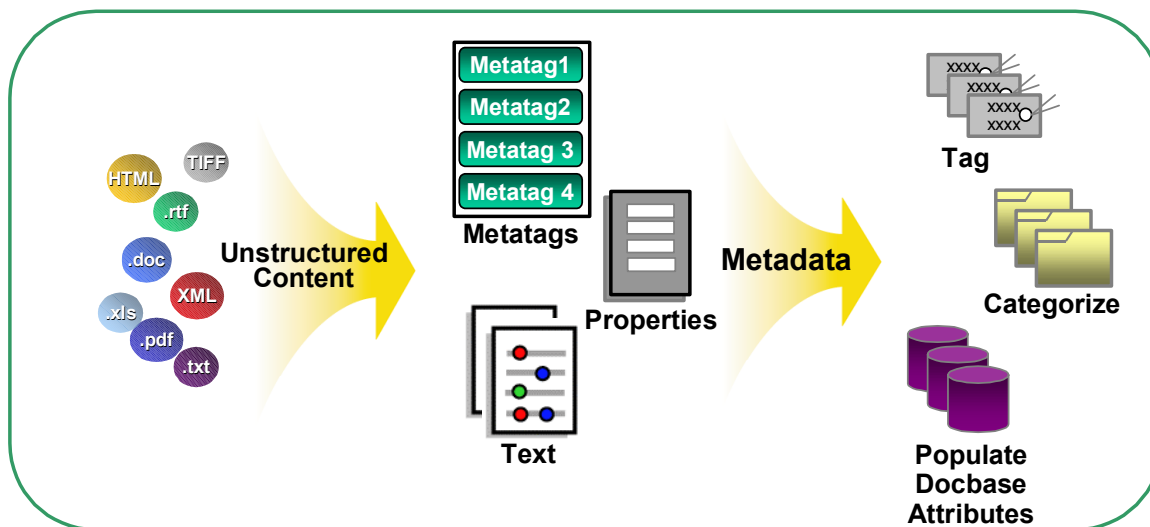
CIS incorporates an **information extraction engine** that extracts file properties, metatags, and delineable text (based on domain map evidence) from the parsed content. This engine generates “actionable” metadata that is used to set content attributes in the Documentum repository. The different sources that CIS can extract values from include:

- **File properties**, such as those in the File/Properties menu in Microsoft Office documents and PDF files
- **Metatags from within HTML/XML documents**. For example:

```

<Title> Content Intelligence Services </Title>
<META NAME="Author" CONTENT="Yue-Jing Lee, Victor Spivak">
<META NAME="Keywords" CONTENT="CIS, Documentum, Personalization">
<META NAME="Copyright" CONTENT="2002 Documentum, Inc.">
<META NAME="Doctype" CONTENT="White Paper, Technical">
  
```

- **Delineable text from within the document structure.** For example, if you know that the second line in a press release announcement always contains the subject of the announcement, you can set up a filter in CIS to extract the value and map it into the Documentum “Subject” attribute field.



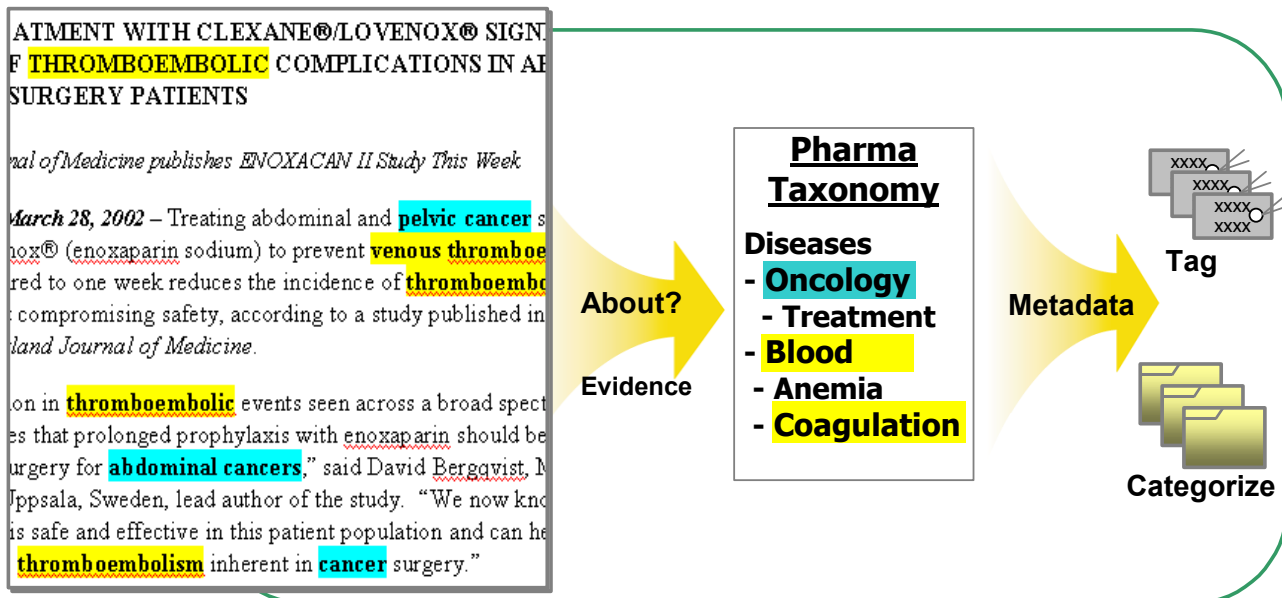
Information Extraction generates actionable metadata and text to enable rich automatic tagging and categorization

### Conceptual Classification

The **conceptual classification engine** is used to determine the “aboutness” of content so that it can be automatically tagged with the appropriate metadata and automatically categorized into the desired repository folder.

The diagram below shows an illustration of how the conceptual classification engine would be used in a pharmaceutical company. For example, an industry-specific pharmaceutical taxonomy is used to define the concepts in the domain map. For simplicity, let’s assume that the values of these concepts are the same as the names of the taxonomy nodes. The conceptual classification engine analyzes the body of the document (which happens to be a press release on a new drug to treat pelvic cancers), and extracts evidence that the document belongs in the Diseases, Oncology, Blood and Coagulation folders. In this particular case, additional concepts and evidence defined in the domain map allows CIS to identify related concepts that are not necessarily an exact match with the taxonomy.

The conceptual classification engine also expands “searchability” and association by categorizing the original content into the main folder (Oncology) and automatically linking it to several other folders (Blood, Coagulation). These related folder names become embedded in the content metadata as keywords upon the completion of CIS processing.

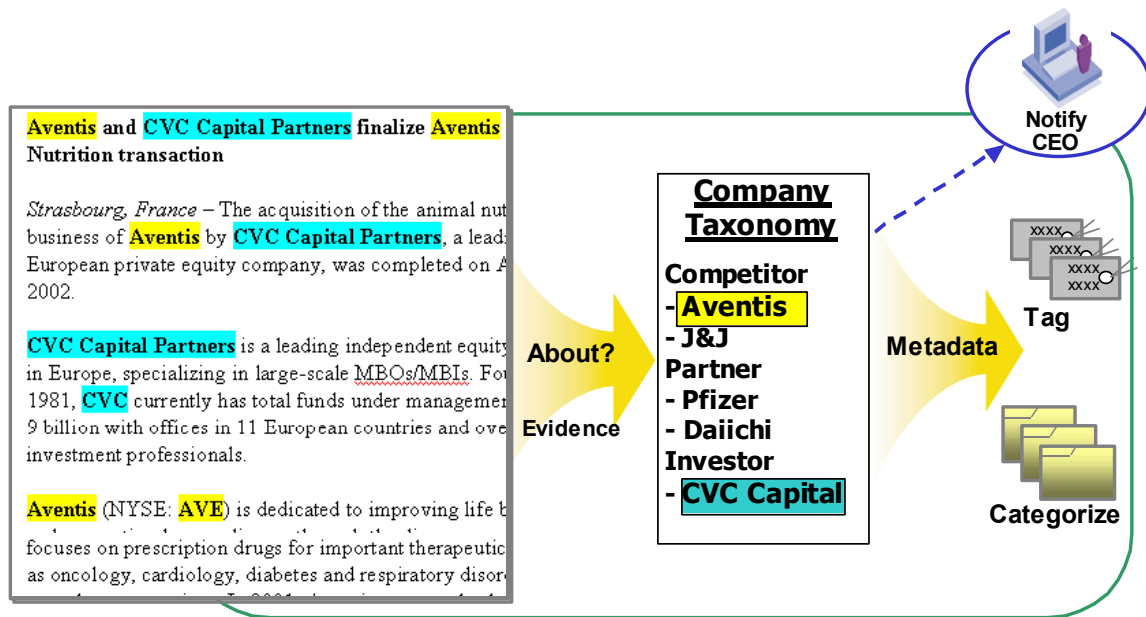


Conceptual classification determines what content is “about” to enable richer associations and expand searchability.

### Semantic Analysis

The **semantic analysis engine** takes conceptual classification to a higher level by incorporating your company’s specific business objectives into the tagging and categorization process, thus driving additional business processes and enabling even richer associations between content. The diagram below includes an example of how competitive intelligence can be enabled by analyzing a recent press release and classifying it according to a company-specific taxonomy. The company-specific taxonomy may include a list of competitors, partners, investors, products, and customers.

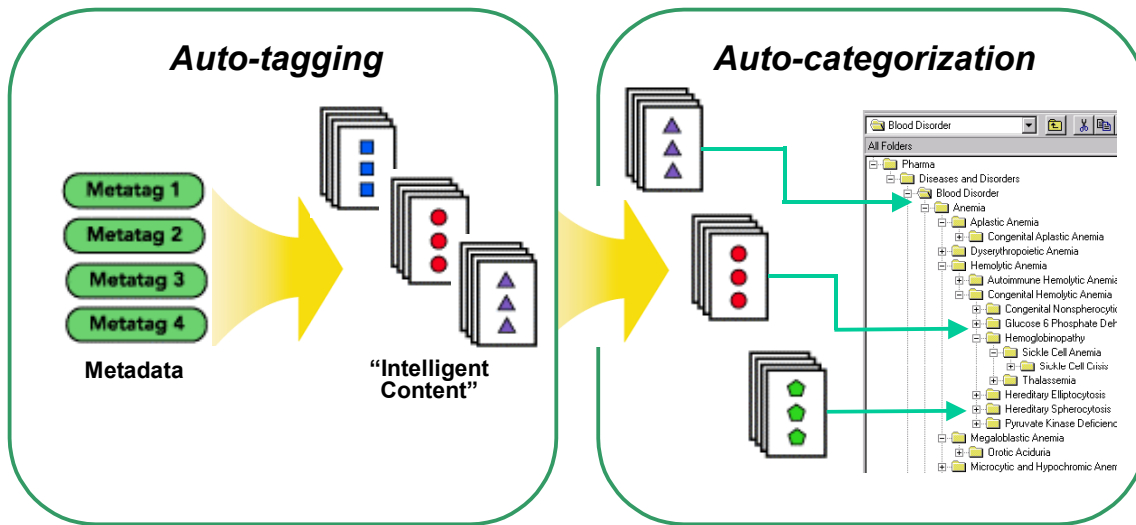
When the press release is processed by the semantic analysis engine, it is automatically tagged and categorized as being related to a “competitor” and an “investor.” A subsequent business process can then be invoked via Documentum workflow to notify the CEO via email or on his PDA that a significant event that may affect his company’s position in the market has just occurred. Now the CEO can take immediate action and respond to this intelligence to ensure his employees are aware of the competitive information and able to respond to comments and questions from customers, analysts, or the press.



Semantic analysis automatically analyzes content to reflect business associations and drive business processes.

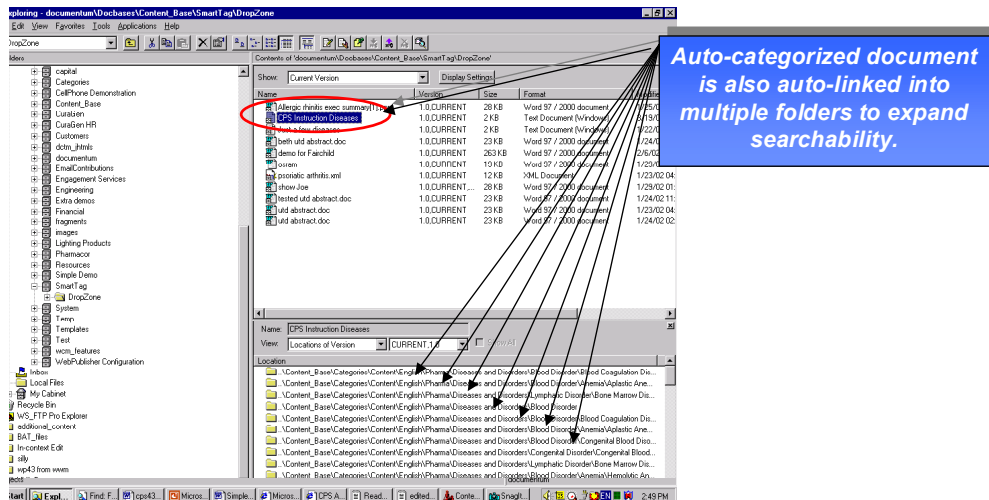
### Tagging and Categorization

Working together, the CIS components — information extraction, conceptual classification, and semantic analysis — enable powerful auto-tagging and auto-categorization capabilities. Through auto-tagging, CIS determines and sets attributes in the Documentum repository so these attributes can be leveraged by search engines and personalization servers. Auto-categorization allows CIS to organize each individual document into the appropriate folder and cabinet based on its contents, enabling users to find content by navigating through a logical folder/file structure. In brief, information extraction, conceptual classification, and semantic analysis make it possible for CIS to auto-tag and auto-categorize content, which in turn enables more accurate searching, easier navigation, and more effective personalization.



Auto-tagging and auto-categorization are the end result.

As mentioned above in the discussion of conceptual classification, CIS not only categorizes the content by placing into the appropriate folder, it also automatically links it to other related folders based on concepts defined in the domain map. This provides for richer associations between content and powerful searching.



CIS auto-categorization links content into multiple repository folders, where links point to the original content.

### Security and Access Control

While CIS makes information easy to find, the Documentum ECM platform takes care of another important issue many organizations face: security. Built-in security capabilities allow you to control access to sensitive content even as you make that content more widely available. While auto-categorization places a given piece of content into a certain folder based on analysis of the information it contains, Documentum security may limit a user's access to

that folder or some of the content in that folder. With Documentum, an administrator can determine which categories a user has access to and even which documents are accessible within a given category.

### Pre-built Starter Taxonomies

So far we've covered taxonomies from an end-user perspective. However, the process of actually building a custom taxonomy is often time- and resource-intensive and can take four to six weeks or longer, depending on how many taxonomy nodes are required. Taxonomy creation often requires special skills such as domain experts, business analysts, and taxonomy scientists.

To help speed time-to-solution and reduce deployment costs, CIS provides several pre-built departmental and industry vertical starter taxonomies based on industry-specific thesauruses such as MESH, CRISP, MedRA, and IEEEE. The pre-built taxonomies are provided in XML format along with documentation to describe their contents. Examples of these pre-built taxonomies include: life sciences, financial services and energy (oil and gas), HR, IT, sales and marketing, customer support, and legal.

#### **Pre-built Taxonomies**

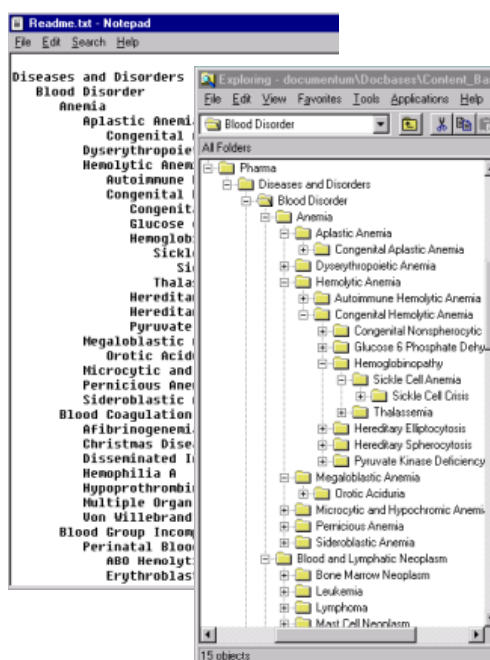
- Provides departmental and industry starting points to shorten deployment
- Examples include sales & marketing, CRM, IT, legal, HR, life sciences, financial services, energy

#### **XML Taxonomy Import**

- Imports taxonomy and automatically generates Docbase folder structure
- Based on industry standards (XML)

#### **Taxonomy Editing**

- Taxonomy creation and maintenance to incorporate new business environments



Content Intelligence Services incorporates rich taxonomy and metadata management capabilities, including pre-built starter taxonomies, XML taxonomy import, and taxonomy editing.

### Taxonomy Exchange Format (TEF)

For standardizing taxonomy exchanges between metadata-generating vendors such as Semio, Autonomy, Verity, Inxight, and personalization servers and categorical navigation through enterprise portals, CIS incorporates a utility for taxonomy exchange format (TEF). The TEF utility converts taxonomies into a standard XML format, automatically generates the appropriate repository folder structure to mirror the taxonomy, maps concepts in the pre-built starter taxonomies to repository folders and, finally, allows for importing of the pre-built starter taxonomies included with CIS.



## Application Examples

### Personalized Web Sites and Cross-Selling

As more and more companies implement Web content management solutions, the need for personalization and content targeting becomes critical in both their customer-facing and employee-facing Web sites. Companies need to deliver the right information, in the right context, and on demand to keep customers coming back for repeat business — and personalized content allows companies to meet that goal.

**Scenario:** An outdoor equipment e-retailer currently has two different classifications of content on its site. One of the classifications contains information about product categories and is used as a product catalog for merchandising. The other classification contains information about the activities for which products are targeted. The activity classification is used mainly by the company’s “learn and share” site, which allows customers to search for and share advice about certain activities.

**Problem:** On the site, there are no relationships established between the two classifications. In order for the e-retailer to target relevant product recommendations to a consumer of an outdoor activity article, the author of the article will need to physically create the links to those product information pages.

**Solution:** Using CIS, these two classifications (activity and products) will be incorporated into a taxonomy so that when CIS processes a “learn and share” article on a specific activity, it will not only classify the article into the correct activity but also pick up the relevant products mentioned in the article. For example, an author creates an article on backpacking for beginners, in which she talks about the precautions that the reader should take when going camping or hiking and suggests a list of tools and gear the reader should bring on the trip. During CIS processing, the article is classified into the “camping/hiking” activity category. In addition, because CIS understands what the recommended camping gear is and stores this information as attributes for the article in the Documentum repository, the Web site will automatically make personalized product recommendations, with product information dynamically linked and displayed directly on the article page. The e-retailer suddenly realizes new selling opportunities while relieving authors of the tedious (and often inconsistent) work of manually searching for and creating appropriate product links.

### Pharmaceutical Clinical Trial Sites

**Scenario:** To help with conducting clinical drug trials, many pharmaceutical companies set up separate Web sites to manage the dissemination and accessibility of information by physicians and patients participating in the trial. In many cases, information accessible by either physicians or patients is segmented. For example, the information presented to physicians may include medical research articles about drugs or illnesses, whereas patients may receive advice about what precautions to take. In addition, this information can often be grouped into different categories (indications, therapies, drug classes) to which the physicians and patients can subscribe.

**Problem:** To support segmentation and subscription on the Web site, content must be categorized. Many times, this requires the author to define or select the targeted viewing audience and the category into which the content should be categorized. This work is tedious and does not allow for categories to be redefined without a lot of work involved in manually re-categorizing existing content.

**Solution:** With a defined taxonomy, CIS can automatically classify content into various categories such as indications, therapies, and drug classes, which can then be used by the site for subscription. In addition, for those categories that share user segments, CIS can be used to further segment the technical documents from those targeted for general audience, so they are served to the appropriate recipient. Finally, when categories for the content change, alterations can be made to the CIS domain map to reflect the new category structure. Once the domain map reflects the desired new categories, re-categorizing of existing content is automated through CIS processing.



## Business News and Event Tracking

With the rapid growth of information on the Internet today, many companies need a mechanism to capture and aggregate important business news that is pertinent only to their industry and to the companies they do business with. CIS can automatically tag and categorize aggregated content, which can then be automatically distributed through Documentum Content Distribution Services (CDS) to a community of internal or external subscribers. Because CIS automatically categorizes content according to a taxonomy structure, it can expose the taxonomy to content subscribers who in turn can select only the categories of information they wish to receive.

**Scenario:** Many investment banks and financial services firms subscribe to news services (Associated Press, Reuters, Wall Street Journal, Bloomberg) to provide them with business information essential to the successful management of their fund portfolio or their customers' assets.

**Problem:** With the amount of information and the number of companies in the market to be tracked, these financial services firms need a tool to automate the categorization of news events by the type of announcement and by companies.

**Solution:** Using CIS and a business event tracking taxonomy, financial services firms can now categorize news into events such as resignations, executive appointments, share buybacks, profit warnings, and earnings announcements, as well as by company and by industry. Now, the fund managers, investment analysts, and industry analysts can subscribe to the pre-categorized news events to facilitate tracking of their portfolio companies.

## Content/Component Reuse Support

**Scenario:** Many companies produce sales and marketing materials such as product briefs, technical data sheets, white papers, and presentations for their products. Often, these documents have common components such as product description, features and functionality, and specifications, which can be shared and reused.

**Problem:** Companies have legacy monolithic documents (such as maintenance manuals, product catalogs, specifications, and standard operating procedures) that contain information that can often be shared and reused in other related documents. These monolithic documents need to be transformed into reusable components that can be assembled as needed. Once these documents are broken down into components, however, it becomes difficult for users to locate the relevant pieces.

**Solution:** The native XML capability in the Documentum repository can be used to “chunk” these monolithic documents into smaller reusable components. Once the documents are componentized and stored as virtual documents, CIS can be used to scan, extract, attribute and categorize each component. Each component will have a unique set of metadata to enable easy search, retrieval, and personalized delivery, allowing users to quickly reassemble standard, approved text and images to create or update documents.

## Business Benefits

Documentum Content Intelligence Services is a core requirement for managing multiple content initiatives and unlocking the value of enterprise content. CIS delivers proven ROI by reducing costs through improved business processes and enhancing productivity — meeting those objectives while increasing customer satisfaction. CIS provides the following tangible benefits:

### Cost reduction:

- Automates manual, error-prone content tagging and categorization methods
- Reduces deployment time through the use of pre-built starter taxonomies

### Enhanced productivity:

- Allows your employees to find what they are looking for more quickly through accurate searches and easier navigation

**Increased customer satisfaction:**

- Promotes higher success rates (and less frustration) in finding information , allowing employees to speed sales cycles, provide enhanced customer support, reduce development cycles
- Enables personalized online interactions with customers by delivering the right content, at the right time, in the right context, on demand

**About Documentum**

Documentum provides enterprise content management (ECM) solutions that enable organizations to unite teams, content, and associated business processes. Documentum's integrated set of content, compliance, and collaboration solutions support the way people work, from initial discussion and planning through design, production, marketing, sales, service, and corporate administration. With a single platform, Documentum enables people to collaboratively create, manage, deliver, and archive the content that drives business operations, from documents and discussions to e-mail, Web pages, records, and rich media. The Documentum platform makes it possible for companies to distribute all of this content in multiple languages, across internal and external systems, applications, and user communities. As a result, Documentum customers, which include thousands of the world's most successful organizations, harness corporate knowledge, accelerate time to market, increase customer satisfaction, enhance supply chain efficiencies, and reduce operating costs, improving their overall competitive advantage.

For more information about Documentum, visit [www.documentum.com](http://www.documentum.com) or call **800.607.9546** (outside the U.S.: +1.925.600.6754).

**Documentum, Inc.**

6801 Koll Center Parkway  
Pleasanton, CA 94566-7047  
phone (925) 600-6800  
fax (925) 600-6850

[www.documentum.com](http://www.documentum.com)

© 2003 Documentum, Inc. All rights reserved. Documentum, and the corporate logo are trademarks or registered trademarks of Documentum, Inc. in the United States and throughout the world. All other company and product names are used for identification purposes only and may be trademarks of their respective owners. Documentum cannot guarantee completion of any future products or product features mentioned in this document, and no reliance should be placed on their availability. Printed in the U.S.A. 60270703V3