Communications of the IIMA

Volume 4 | Issue 1 Article 9

2004

Enterprise Content Management: Understanding the Taxonomy

Jake Zhu
California State University, San Bernardino

Tony Coulson
California State University, San Bernardino

Miyuan Shan
Hunan University

C.E. Tapie Rohm Jr.
California State University, San Bernardino

Enterprise Content Management: Understanding the Taxonomy

Jake Zhu

California State University, 5500 University Parkway, San Bernardino, CA 92407
(909) 880-5768, Fax: (909) 880-7176, jzhu@csusb.edu

Tony Coulson

California State University, 5500 University Parkway, San Bernardino, CA 92407
909-880-5768, Fax: (909) 880-7176, tcoulson@csusb.edu

Miyuan Shan

School of Business Administration, Hunan University, Changsha, Hunan, P. R. China, 410082
+86 (731) 882-3903, Fax: 86 (731) 882-3670, shanmiyuan@163.com

C.E. Tapie Rohm, Jr.

California State University, 5500 University Parkway, San Bernardino, CA 92407
909-880-5786, Fax: (909) 880-7176, trohm@csusb.edu

ABSTRACT

In an effort to understand content, its taxonomy and management, we define content management as the process of managing the creation, categorization, delivery and archiving of content in which workflow and business processes are maximized. Built upon what industry experts have to offer about enterprise content taxonomy, we have enacted a framework to better understand, create and manage enterprise taxonomies. This framework will be able to integrate rich content, modern computing technology, advanced analytics, automated workflow, business requirements and human resources into a process where companies can easily manage their information for decision making and gaining competitive advantages.

INTRODUCTION

The rapid developments of the Internet technologies and the needs for seamless content delivery have shaped a favorable environment for the content management (CM) market. Through much of the nineties of the last century, information technology (IT) has created myriads of new business opportunities for entrepreneurs and organizations to cash in the Dot-com phenomenon. However, technology alone cannot sustain competitive advantage. The slow-down of the economy since the turn of this century has taught businesses hard-to-forget lessons that the gold rush on the Internet is over, and that they have to search for more solid business strategies to gain business advantages and remain competitive. They turn to content management (McKeever, 2003; Pachet, 2003; Sumner, 2004; Trappey and Trappey, 2004; Upshall, 2003).

Although interest in enterprise content management (ECM) practice and related technology advances has created an exciting array of demand for CM systems by organizations as well as e-businesses, Content management is not yet well understood by the research communities. Much of its publicity has been generated by industry experts who echo their voices and try to promote certain CM technologies and case based management practices through trade magazines and the Internet media. Enterprise content management is a new market formed by the convergence of several existing markets including the traditional integrated document management (IDM) market, the media asset management (MAM) market, and the web content management market (Zimmer, 2001). Content management plays an important role in an e-business because it enables the seamless flow of product information between suppliers, manufacturers, distributors and customers (Trappey and Trappey, 2004). With the proliferation of web content and the growing demand of customers for web-based services, companies have an increasing need to support their web content with web content management (WCM) systems (McKeever, 2003).

Central to ECM is the concept and practice of an enterprise content taxonomy which provides a shared vocabulary for companies to classify and organize documents (Venkata, 2002). What are some of the major issues in content management? What is the best way to categorize content? How do CM systems offer personalized information for decision making? How does a CM application support business processes by workflow automation? These questions are all related, in some way, to an important aspect of content management – enterprise taxonomy. The purpose of this article is then to offer understanding of enterprise content, its classification, and categorization by analyzing what industry experts have said about content taxonomy. In so doing, a complete framework is formed for interpreting content taxonomy, classification, categorization and its management.

But what is content management? In the following section before we illustrate the methodology for data collection and analysis, we will offer a view on this.

CONTENT MANAGEMENT DEFINED

Content management implementation provides end-to-end CM for different types of content, from creator to ultimate consumer (Rudy, 2001). According to him, business-critical content

management applications require integration with authoring tools, version and workflow management, translation to web formats, web page layout, and distribution through standard, mobile, and wireless browsers.

In order to understand CM, one needs to define the variables that relate to the concept: content and management. Content involves its creation, categorization, sorting, delivery and archiving. Rudy (2001) categorizes business content into three types:

- *Unstructured content*: Unstructured content is characterized by all of the operational, non-database content that drives a business. Typically unstructured content includes text documents, images, spreadsheets, presentation materials, and may also include drawings, reports, email, video and audio.
- *Structured content*: Generally considered the domain of the database systems.
- *Web content*: Unstructured and structured content that exists independently of a web site.

Management involves linking enterprise content to business workflow and processes. A workflow system integrates process, content, people and systems—inside and outside the firewall (Donnelly, 2002). According to Lamont (2002), workflow has evolved to encompass a broader discipline referred to as business process management (BPM), which includes in the flow both people and enterprise applications. These applications not only route information or documents to individuals for decision-making, but also automate decisions by linking business rules to data sources. Much of the content that organizations need to provide to their customers, employees, and business partners is part of a business process (Christian, 2001). Therefore, an organization's content, business processes and workflow should be well integrated together to maximize information value.

To summarize, content management is the process of managing the creation, categorization, delivery and archiving of content in which workflow and business processes are maximized.

METHODS

DATA COLLECTION

In order to obtain data for analysis, we went through a two-step identification process. First, we employed an extensive search on the Internet for publications that extensively circulate content management (CM) materials to educate information management professionals and potential customers in which we identified three industry trade magazines. Table 1 is a summary of the three trade magazines and their missions or purpose statements.

The three publications were chosen according to the following criteria: 1) major contribution to the understanding of ECM in terms of knowledge and education; 2) extensive coverage of CM technologies and best CM practices; 3) major expert support and decision making awareness for businesses and organizations in the CM field; 4) regular publication with Internet access for around 4 years already which is the period when CM technology and practices have flourished.

Of the three trade publications, CMSWatch and E-Doc Magazine explicitly stated that their mission is to promote understanding and awareness of CM or ECM which we will use interchangeably throughout this article. KMworld was chosen because of its major contribution to the understanding of ECM and its extensive coverage of best CM practices even though its mission uses KM instead of CM. A "content management" keyword search on its web site yielded over 300 articles that relate to CM and its technologies (Search was performed on March 31st, 2004). This was exclusive of news and product advertisement related to ECM. That is about over 50 articles on CM per year in its more than 6-year publication history on knowledge management. Besides the regular publication of KM and CM materials under KMWorld Magazine, it also has KMWorld Special Publications that publishes best practice in special fields such as ECM or enterprise portals as supplements to the regular issues.

Publication Title	Mission
KMWorld *www.kmworld.com*	KMWorld is the leading information provider serving the Knowledge Management systems market. We inform our more than 56,000 subscribers about the components and processes - and subsequent success stories - that together offer solutions for improving your business performance
CMSWatch *www.cmswatch.com*	The goal of CMSWatch is to educate and inform users about current and emerging content management trends, and to empower them to identify, evaluate and select content management solutions that are appropriate for their needs. We seek to be a resource for the web development community and to contribute to the growing body of knowledge about content management systems.
E-Doc Magazine *www.edocmagazine.com*	E-DOC Magazine is an unrivalled and unbiased industry source for information on enterprise content management (ECM) topics that help management professionals worldwide make informed and effective technology decisions.

Table 1: Publication and Mission:

Second, we use "taxonomy" as the keyword to search each of the 3 publications sorted by relevance. We identified 12 articles from the 3 publications which seem to contribute ultimately to the understanding of content and its taxonomy. Table 2 is a presentation of the 12 articles:

These articles were chosen because of their relevance to the theme of this article which is to understand content taxonomy. Specifically, we used the following criteria when we chose an article 1) contribution to the understanding of ECM; 2) provide knowledge or insights on

enterprise content taxonomy, content classification and categorization; 3) build links between technology, people involvement, information workflow and business processes.

Table 2 presented all the articles identified. A unique identification number was assigned to each article. The letter K, E or C in the ID section of the table is the initial letter of the publication, and the numeric code identifies each article in the same publication sorted by date. It has no weight on the content of the article itself. These IDs will be used later in the result section for citation purposes.

Publication Title	ID	Article Title	Publication Date
KMWorld	K1	Building an Enterprise Taxonomy	November, 2001
	K2	Taxonomies, Categorization and Organizational Agility	October, 2002
	K3	Leveraging Unstructured Information to Enhance Business Intelligence	January, 2003
	K4	Content Taxonomy—Reduce Your Exposure	September, 2003
E-Doc	E1	Content Categorization	January, 2002
Magazine	E2	Taxonomy: More than a topic tree	January, 2002
	E3	Automatic Categorization, Taxonomies, and the World of Information: Can't Live With Them, Can't Live Without Them	November, 2002
	E4	Auto-Categorization and Records Management	March, 2004
	E5	The Value of Indexing	March, 2004
CMSWatch	C1	The Case for Personal Web Publishing	November, 2001
	C3	Value of Organized Knowledge	January, 2002
	C4	A Metadata Primer	February, 2003

Table 2: Article List:

DATA ANALYSIS

The data analysis for this article involved a four-step process. First we tried to identify what is important in content taxonomy – the key concepts. These concepts later became categories in the analysis. Then we generated properties for these conceptual categories. Properties are unique characteristics of these categories. As Glaser and Strauss (1967) put it, a property is a conceptual

aspect or element of a category (p. 36). Next we compared different perspectives within the same category. We finally form our own framework based on the above analysis in the discussion section – theory generation.

The data analysis process is much like a grounded theory approach (Glaser and Strauss, 1967; Locke, 2001) where incidents form categories from which theories can be generated. Also like doing grounded theory research, we start our project with no predefined propositions or theoretical vantage point. Our purpose is again to understand content and its taxonomy from the industry perspectives.

The result of the analysis will be presented in the next section.

RESULTS

The search for the meaning of an enterprise content taxonomy from data yielded three major conceptual categories:

- Definition of Taxonomy
- Methodology for Taxonomy Creation
- Managing Taxonomy

Table 3 is a presentation of all three conceptual categories, their properties and their related article IDs:

Category	Property	Article ID
Definition of Taxonomy	• Content tree directory • Classification rules • Function: o Tools for information management o A business process	K1, K3, K2, K4, E2, C2, C1
Methodology for Taxonomy Creation	• Manual/Automation/Hybrid • Content Driven • Process Driven • Technique Driven	K1, K2, K3, E1, E2, E3, E4, E5, C1, C3
Managing Taxonomy	• Taxonomy creation/Content inclusion/Interface Creation/Population/Maintenance • Total Taxonomy Lifecycle Management	K3, E1, E2

Table 3: Categories, Properties and Articles: