Content Management-The Magic behind the Web

JoAnn T. Hackos, President, Comtech Services, Inc.

"Where is that white paper we need for the new product Web page? Isn’t Jan writing it? Has she finished? Was Tom supposed to approve it? Did Jan get it to him already? We’re supposed to have it up on the site by Friday at 5 pm. I guess I’ll have to call Jan to find out if she’ll make the deadline. I know Tom is on the road; I wonder if he’s reading his email?"

If this scenario sounds familiar, you are not alone. Managing even the simplest content has inherent problems. Whether you have tens or thousands of employees, content doesn’t happen by accident. Unfortunately, many developers of Web content-management systems seem to believe that content has no life cycle before it appears at the door of the Web manager. Despite all assumptions to the contrary, content has a long and sometimes convoluted life before it is ready to go live. But in too many organizations, that life cycle starts on someone’s hard drive and remains there under little or no management. In a typical case, content begins with Jan, who gets the assignment to prepare a white paper. She’s not the subject-matter expert (SME) on third-tier widgets, so she has to begin by finding those who are and getting copies of whatever information they have on their hard drives or departmental servers. Most of the time, those nuggets of source material are transferred by the SMEs to the content developer by email. Now there are two copies of the source material, one on the SMEs’ drives and another in the document that Jan is preparing. If the SMEs update any of the source information, it will be sheer luck if Jan finds out. The success of an informal information flow depends upon having conscientious people with extra time on their hands.

**Content Access**

With content management, as soon as the white paper is in the planning stages, the source material on the SMEs’ hard drives is checked into the content database with appropriate labels to categorize the content that trace it back to the originators. The categorization of content (using metadata dimensions) follows a scheme that promotes the interests and concerns of the business. The categorization scheme might include administrative information about the authors as well as categorization that label the information according to brand, product model or version, geography, language, and others.

An author like Jan designates a "virtual" location in a working area once she begins work on the white paper. All the source material is catalogued in the working area while the white paper is under development. Anyone who might have to take over Jan’s responsibilities in an emergency knows exactly where to find all the relevant material without chasing around to different departments, getting access to obscure server areas, or finding people who might know where the sources originated.

But you might argue, can’t we handle source control by having well-labeled work areas, folders, and files without content management? Just what value does content management add to the equation?

Jan’s working area is just one small piece of a much bigger content-management problem. In most organizations, there are thousands of folders holding tens of thousands of files, many of them applicable to more than one ongoing project. Searches that are restricted to file and folder names rarely yield usable results. Finding a particular piece of source material is like finding the proverbial needle in the haystack, but now there’s not a single haystack but hundreds of them.

Content-management systems provide mechanisms for assigning metadata to individual documents or even elements within documents if additional granularity is needed. For known situations, much of the metadata needed to identify the author, subject, date, and other data can be automated or semi-automated so that the author selects from a predetermined and limited set.

Automated content analysis and categorizing tools are also making a strong entry into the content-management market. The purpose of these tools is to analyze and label content using textual and domain analyses, sometimes including elaborate context algorithms. Be aware, however, that automated analysis has its limitations. If the authoring environment is under your control, a well-planned taxonomy with authoring aids may be easier to implement and maintain.

**Workflow**

As Jan works on the white paper, she sends some of the content out for review to the SMEs and to the product managers who are responsible for the final approvals. When her draft is ready, she checks the draft back into the content-manage-
Supporting Version Control

Jan is ready to send the white paper out for final approval before it is published to the company’s Web site. She checks the draft in as complete and the workflow system routes it to the VP of Marketing. The marked up copy Jan gets back indicates that the VP would like to see a stronger ROI analysis in the paper. In a previous version, Jan had included that analysis but had decided to omit it because of comments from the engineering people. She uses the content-management system to trace back to the earlier version where she had added a note that she’d removed the ROI analysis. She then inserts the ROI section into her final copy. She routes it back to the VP, who signs off immediately. Version control saves her time because she will not have to write the ROI analysis again, and it helps her make changes easily.

Version control and check-in/check-out security provides staff members with the flexibility and security they need to manage content during the development life cycle, well before it is deployed on an intranet or Internet site.

Not only does content management support information that will be published to the Web, but it also supports an information-development life cycle that is completely internal to an organization. Content authors are able to secure their information, make information easily available to others in their workflow, and provide access to information that they need to enhance internal business activities.

In any organization, information created in one part of the organization is critical to the business activities of many other parts of the organization. But for the most part, that information is difficult to locate. Typically, content is exchanged through a flood of emails. Behind every email interchange is the time spent by each staff member to request information and fulfill the requests of others.

Calculating the Cost of Content Chaos

How much does this unproductive activity cost? More than you would imagine. Conservatively, staff members spend 5 to 10 percent of their time locating content and directing it to others. What is the cost of 5 to 10 percent of people time in your organization? Even if you calculate only the time spent by core staff, you will easily end up with millions of dollars of lost productivity in a mid-size organization.

So is content management in your future? If you don’t think so today, think again.

Content Management Systems

AuthorIT
www.author-it.com

Lightspeed Astoria
www.lightspeed.com

Documentum 4i
www.documentum.com

J D Edwards Enterprise Content Manager
www.jdedwards.com

Live Linx
www.livelinx.com

Noonetime Constellation
www.noonetime.com

Oracle iFS
www.oracle.com

Trisoft InfoShare
www.trisoft.be

XyEnterprise Content@ XML
www.xyenterprise.com

About the author

JoAnn Hackos, PhD, is President of Comtech Services, a content-management and information-design firm based in Denver, which she founded in 1978. In her new book, Content Management for Dynamic Web Delivery, Dr. Hackos explains the content-management strategy that she developed for companies such as Nortel, Motorola, Cisco, and others and walks readers through the stages of effective Web content management. She can be reached at joann.hackos@comtech-serv.com.

Content Management for Dynamic Web Delivery
John Wiley & Sons
Available February 2002
ISBN 0-471-08586-3

For more information about her new book, log on to www.comtech-serv.com.