Understanding Web content management systems: evolution, lifecycle and market

Susan McKeever
School of Computing, Dublin Institute of Technology, Dublin, Ireland

Keywords
Information, Internet, Database management systems, Publishing

Abstract
The volume of digital content available on the World Wide Web has increased dramatically over the past six years. Some form of Web content management (WCM) system is becoming essential for organisations with a significant Web presence as the volume of content continues to proliferate. WCM systems have evolved rapidly from the basic HTML editors of six years ago, to the sophisticated content production and publishing tools available today. This paper presents a WCM hierarchy, examines the underlying Web content management lifecycle, and identifies the key market trends for WCM systems.

1. Introduction

All Web sites display and/or process content in order to convey information to their audiences. In parallel with the growth of the Internet, content volumes have grown rapidly. In 1992, there were just 1,000 pages on the Web. As of June 2000, over two billion Web pages were posted on the Internet (Ektron, 2001). As at December 1996, only 32.7 per cent of Fortune 150 firms had Web sites, according to Cox et al. (2001). This had grown to 94.7 percent by January 1998 (Cox et al., 2001). The current rate of growth shows no signs of dissipating.

Some form of content management (CM) process or system is becoming essential for all organisations with a significant Web presence as the amount of digital content continues to proliferate (Butler Group, 2001). As the Web moves from small, informally designed Web sites into large, rapidly changing sites, the need for strong management tools has become pressing. Software product companies have moved to address this need and call their offerings content management systems. Such systems help a business to set-up and organise their Web site(s), so that the Web sites can grow and change rapidly while maintaining high quality (Boiko, 2001).

In this paper, the author presents two models that enable a clearer understanding of Web content management and its underlying activities. The WCM four layer hierarchy presents a view of WCM as a series of layers, from the Web audience layer through to the content level. The WCM lifecycle then discusses the iterative nature of WCM and the core components of the lifecycle.

The research for this paper was obtained from industry publications, industry reports, journal sources, vendor commentaries, and online articles published by respected figures within the computer industry.

The paper is structured as follows: Section 2 examines the evolution of Web content management systems. Section 3 introduces the WCM four layer hierarchy. Section 4 discusses the constituent phases of the content management lifecycle. Section 5 examines the market for WCM systems, identifying the key players and future market trends. In the concluding section, the author summarises the new models presented in the paper, and suggests several research extensions.

2. Evolution of Web content management systems

The era of the Webmaster as Lone Ranger is over. With technology and business requirements evolving at “Web speed”, the time has passed when one guru could set up and manage a Web site to provide a corporate presence on the Internet (Seeley, 2000).

Early Web sites

When the WWW was in its infancy, Web sites largely consisted of static hand crafted HTML code on convenient text editors (see Figure 1). WCM “systems” consisted of basic scripting tools for assisting in the production of HTML. The HTML pages were either edited directly on the live site or were deployed via file transfer protocol (FTP) directly to a live Web site (Nakano, 2001).

Pages largely consisted of static text, links and a limited number of graphical images. A number of common problems tended to occur: poorly coded HTML code, broken tables, disconnected links, poor quality content and missing graphics. Responsibility and capability for managing the site lay with the Web site manager or “Webmaster”. The
actual content authors within the business were completely reliant on the technical skills of the Webmaster in order to deploy their content. However, at this point in time, having a business Web site was considered good practice in itself, even if the content was not of a high standard (Hoffmann, 2000).

Gradually, standards of deployment improved, and it became common practice to use a separate server from the live site for development. Procedures for, and automation of site update and deployment tasks became more commonplace. There was a continued high reliance on the Web master to manage and administer the Web site(s) (Hoffmann, 2000).

**Increasing maturity of WCM**

In the late 1990s, Web site complexity exploded along multiple dimensions: higher volumes of content, more dynamic content, higher numbers of visitors, and increasingly complex supporting hardware and software. Dynamic content of all kinds is now assembled from back-end databases and middle tier application servers, coded in multiple scripts and languages, and served up via Web server farms distributed across the a single country or internationally. The number of hits has increased exponentially, and the content volume and variety of content types has risen (Hoffmann, 2000). The responsibility for content deployment has partially or completely moved from the Webmaster and his/her team out to the actual authors within the business, enabled by simple WCM system interfaces for content contribution (eGrail, 2001). Web site errors are no longer an acceptable occurrence for Web site visitors. A broken link or out of date content on a company Web site is now not just an embarrassment but can impact reputation and business profitability (Hoffmann, 2000).

Figure 2 contains a diagram of the progression of content management systems, illustrating how content management has developed over the past six years.

Having introduced the evolution of WCM systems, Section 3 introduces a simple hierarchy for understanding the full context of modern WCM.

### 3. Web content management: a definition and hierarchy

Content is the feeder mechanism for all business processes. And always has been (Moore, 2001).

One of the core components of Web applications is the provision and management of content. Types of content include, but are not limited to, Hyper Text Mark-up Language (HTML), Extensible Mark-up Language (XML), images, videos, documents and dynamic content generated from relational databases (Frost and Sullivan, 2001). Other types that may form part of a Web application are e-mails, catalogues, technical documents, audio and video, databases reports and e-commerce transactions (Butler Group, 2001).

Boiko (2001), author of “Understanding content management” provides the following definition of content management:

At the highest level, Content management is the process behind matching what “you” have with what “they” want. “You” are an organisation with information and functionality of value. “They” are a set of definable audiences who want that value (staff, partners, customers). Content management is not just a way to create large Web sites, but upon closer examination, it is
in fact an overall process for collecting, managing and publishing content to any outlet.

Content management, according to Boiko (2001), ranges from content authoring or acquisition through to content publishing/deployment and Web site management. In turn, the audiences for a company’s Web content can be identified as staff, via intranet, and partners, suppliers or customers via an extranet site and/or Internet site (Ashenden et al., 2001).

The full context of WCM has illustrated graphically using the WCM four layer hierarchy as shown in Figure 3. The hierarchy consists of four layers, reflecting each of the interacting layers in content management: content, activity, outlet and audience. (In fact, a fifth “device” layer containing information such as personal computer, mobile phone and hand held computer is relevant to the hierarchy, but is outside the scope of this paper.)

Layer one of the hierarchy, the content layer, consists of the types of content that may require management by a WCM system or process. The activity layer then represents at a high level the activities involved in managing the content layer: content within the content layer is created, and then deployed into an appropriate outlet. Moving up through the hierarchy, the outlet layer defines the types of Web outlets through which content can be accessed: intranet, extranet and Internet. The audience layer then defines the groups of people who will be interacting with the Web outlets or presence(s) of a company.

In tandem with the WCM four layer hierarchy, the following definition of Web content management is provided by the author:

Web content management incorporates the activities involved in the creation and deployment of digital content to Web based audiences, where these audiences may consist of customers, suppliers, partners and staff accessing Web content via extranet, Internet, or intranet. A WCM system consists of the software tool(s) used to provide automated support of WCM activities.

In Section 4, the author defines the lifecycle of Web content management within an
organisation and discusses the phases or steps that constitute this lifecycle.

4. WCM lifecycle

Understanding the process or lifecycle of Web content management is fundamental to understanding the concept of WCM. In researching the process of managing Web content within an organisation, a generic, iterative lifecycle emerges based on four steps or phases. The research sources (Boiko, 2001; eGrail, 2001; CMS, 2001; Byrne, 2001; Merant, 2001) use differing terminology for the steps within the process or lifecycle, but generally concur on the actual content of the phases involved. These research sources imply a sequential set of phases, leading to a linear WCM process. In the view of the author, however, WCM within an organisation is a continuous series of iterative phases of creation and deployment, reflecting the continually changing and dynamic nature of the content that an organisation will display. This concurs with Hong and Lee’s (2002) iterative knowledge management lifecycle whereby knowledge is continuously captured, developed, shared and used. The iterative phases of content creation and deployment are governed and supported by control and administration, and workflow phases. This is illustrated in the lifecycle diagram in Figure 2.

The lifecycle contains two iterative phases that will continue as long as the organisation continues to have Web presence(s): the collection of content, and the delivery or publishing of that content on the Web. Collection, as termed by Boiko (2001), also called “authoring” or “creation” (CMS, 2001), refers to creating content or acquiring content from an existing source. The process of content creation may involve many people within a department or throughout a company (Byrne, 2001). Each person may have a distinct role, such as editing or reading documents within a specific department or function. Typically, users will be non-technical users (eGrail, 2001) who will require easy to use, standard interfaces. Content collection may involve large numbers of users distributed across separate locations (Microsoft, 2001), so complexity of implementation will be reduced if the system has a thin client. Indeed, according to Merant (2001), non-browser client software will prove to be a significant barrier to the use of the system.

The range of possible functions in a WCM system to support content collection, as researched by the author, includes:

- Standard tools, with standard user interfaces, for content creation (PWC, 2001).
- Multi-user support and distributed authorship to enable geographically or departmentally distributed content contributors (PWC, 2001).
- Separation of content from presentation, centralised definition of look and feel (Butler Group, 2001).
- Support for content syndication (Merant, 2001).
- Content versioning for authors (Merant, 2001).
- Support for relevant content types (Hoffmann, 2000).
- Database forms support for catalogue type data (Merant, 2001).
- Support for localisation/multiple languages (Herbig and Palumbo, 1998).
- Shared database for content storage (PWC, 2001).
- Thin client for easy rollout (eGrail, 2001).
- Real time access to CM functions (Day, 2001).

Delivery of the content to the Web site(s), also called publishing (Boiko, 2001), deployment (Nakano, 2001) or distribution (Byrne, 2001), involves making the content available to Web users by extracting components out of the content repository and constructing specific Web site pages. In a non-automated environment, this is tedious and time consuming. With the support of a content management system, Web pages can be constructed dynamically from the content repository or if appropriate, may consist of static pages constructed from HTML. Static pages are more appropriate where the content of the page is not variable – for example, the text on a Web site providing a general background to the company (Boiko, 2001).

Content delivery has several other dimensions that need to be considered: the facility to control versions of content, the ability to roll back a Web site, or subset of a Web site to a previous version, the management of separate locations or environments for development content versus production content and the ability to delivery content to multiple channels.

The research sources used for this paper identified the following WCM system features to support content delivery:

- Support for dynamic content (Boiko, 2001).
- Automatic Web page link checking (Seeley, 2000).
• Intelligent Web data error checking, such as product price sense checking (Seeley, 2000).
• Separate environments for development and production (Merant, 2001).
• Content version rollback facilities (Merant, 2001).
• Multi-channel support with re-use of content for multiple end user devices (Steiner, 2001).
• Scheduling of automated site changes (PWC, 2001).
• Content personalisation features (Koong et al., 2002).

Supporting both content and delivery is the Workflow component of the WCM lifecycle. Workflow enables collaboration on WCM tasks, allowing content to move quickly and efficiently through the development and approval processes into production. This workflow can be single threaded or involve simultaneous tasks, may require active alerts to participants in the process to alert them to actions required, and may require overall monitoring to eliminate bottlenecks in the publishing process (Merant, 2001).

In some organisations, the workflow process may be informal – lacking clearly defined roles, approval steps and expected user response times. An informal or unclear workflow process cannot be supported by a content management system, and requires clarification, and implementation of the clarifications as part of the content management system implementation (Byrne, 2001). An absence of established workflow and approval process can also result in costly and/or embarrassing errors if unapproved content, or incorrectly scheduled content is loaded onto a Web site.

The key system features to support the workflow component include:
• Flexible, multi-threaded workflow (Merant, 2001).
• Workflow monitoring and control features (Seeley, 2000).
• Support for workgroups (Merant, 2001).

The final component of the content management lifecycle is control and administration. The control and administration activities support the remaining phases in the WCM lifecycle. These activities include identification and administration of user roles, identification of groups of users with group level functions (e.g. marketing material approval group), management of the content data repository, categorisation of data within the data repository, verification of security requirements, monitoring of the approval workflow and elimination of workflow bottlenecks (Boiko, 2001).

The WCM system features available to support the control and administration features include:
• Pre-defined or custom user role definition and user security (Merant, 2001).
• Ability to specify metadata to support the information architecture or taxonomy (Koong et al., 2002; Merant, 2001).
• Audit trail facility to capture record of content related activities (Merant, 2001).
• Reporting functions to report on, for example, site visitors, visitor activity, and deployment history (Hoffman, 2000).

This section presented the WCM lifecycle, enabling the reader to understanding the core steps or phases in WCM. In the next section, the market for WCM systems will be discussed, identifying the main players in the market and the key market trends.

5. WCM system market trends

Growth in the WCM system market has been rapid. In 2000 alone, the worldwide content management market experienced growth of 89.8 per cent (Chung, 2001). The content management market is growing twice as fast as customer relationship management and enterprise resources planning systems, over four times greater than worldwide IT market in general and six times faster than the IT hardware market (Strategy Partners, 2001).

According to Ovum (Ashenden et al., 2001), the CM systems available on the market today have evolved from several different sources:
• Document management vendors have re-branded as Web enabled content management tools. Examples are FileNet and Documentum (Ashenden et al., 2001).
• E-commerce and internet infrastructure vendors are becoming content management vendors. Examples are Microsoft’s content management server and IBM’s content manager (Butler Group, 2001).
• Many purely Web focused vendor organisations, such as Interwoven and Vignette, have started up in the last two to six years.
• Single component vendors offering just a single specialist CM function, such as catalogue management, have also established themselves in the market (Strategy Partners, 2001).

The top five WCM vendors own a combined market share of just 28.8 per cent (Strategy Partners, 2001). The market in 2001 was led by
Broadvision, Divine, Documentum, FileNet, IBM, Intervoven, Stellent and Vignette although Microsoft is expected to strengthen its position after buying Canadian company Ncompass (Butler Group, 2001). In 2001, the only company to increase its market share by more than 5 per cent was Intervoven, increasing from 8 per cent to 13 per cent which places it as the top vendor (Frost and Sullivan, 2001).

As the market matures, a number of trends can be identified:

- The worldwide spend on CM software and services is forecast to grow at a compound annualised growth rate of 28.5 per cent up to and including 2005, to reach a total spend of $10,545m (Strategy Partners, 2001).
- Spend by companies on CM systems will be funded by customer-led revenues, as opposed to dot.com funding, with companies carrying out thorough selection processes to ensure value for money (Strategy Partners, 2001). Giga Information Sources estimate that the current average licensing cost for a high-end Web CM system is $500,000, whilst that for midrange is $300,000. However, prices (and breadth of functionality) vary considerably, with licence costs ranging from under $200,000 for basic authoring tools such as Macromedia, to more than $800,000 for market leading products such as Vignette’s content management server (Kontzer, 2001).
- There will be an increase of smaller, niche, vertical solutions as vendors seek ways to differentiate their products (Butler Group, 2001). More specific applications in the areas of e-learning, community system, e-government and cataloguing will emerge over the next year as vendors recast their products into vertical applications (Strategy Partners, 2001).
- Large established software vendors such as IBM and Microsoft are manoeuvring to establish their position through acquisitions and strategic alliances, and will increase their market share as their product offerings widen (Butler Group, 2001).

Given that the smaller CM system suppliers are likely to move towards niche solutions, and the larger established software vendors will grow their share of the CM market, it is reasonable to conclude that the newer vendors of the top tier CM systems will face the greatest challenge in maintaining or growing their market share.

6. Summary/conclusions

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 systems. Current estimates indicate that the market for content management systems is growing at a rate of over 29 per cent per year worldwide.

WCM systems have matured rapidly in the last six years, enabling Web sites to mature from basic HTML text based sites, with Webmaster dependence, to sophisticated, distributed author, multi-tiered architecture sites. The players in the market consist of the previous document management system providers, larger software providers and new, purely Web focused vendors. As the market matures, the larger software vendors such as IBM and Microsoft will expand their product offerings, and an increasing number of niche CM applications will appear.

In this paper, the WCM four layer hierarchy illustrates the four interacting layers in content management: content, activity, outlet and audience. The hierarchy presents a clear view of the scope of WCM. The WCM lifecycle is then defined, based on four main phases: content collection, content delivery, workflow, and control and administration.

For further research, the author suggests that WCM lifecycle presented in this paper can be broken down into smaller constituent phases or steps, in order to enable a deeper understanding of WCM systems’ scope and potential. As an additional area, the impact of WCM systems on organisational business processes would merit further discussion. The availability of easily accessible Web content to a company’s staff, customers, partners and suppliers will impact many business processes within the organisation, such as customer service processes via a customer facing Web site or human resource processes for dealing with staff. Future research could investigate the impact of Web content on these types of business processes, understanding the cost and effectiveness impact of replacing paper or people driven processes with digital content driven processes.

References


Ektron (2001), Effective Web Content Management: Empowering the Business User while IT Maintains Control, Ektron, NH.
Merant (2001), The 10 Critical Elements in Choosing a Content Management Platform, Merant Corporation, Hillsboro, OR.