



Digital Assets Content Management System

White Paper

Prosentient Systems Pty Ltd

Edmund Balnaves

BACKGROUND.....	2
CONTENT MANAGEMENT.....	2
DIGITAL ASSET MANAGEMENT	3
LITERATURE REVIEW	3
REFERENCES	6

BACKGROUND

The effective management of content is a key challenge in the digital era. The publishing of websites is often a spectacular and expensive “once-off” exercise that delivers an impressive result in the short term, but in the long term represents many challenges for content management and re-use. Currency of website content rapidly becomes an issue. Similarly, many organisations have a wide range of text, image, sound and video assets that are distributed on a range of systems, but which are poorly management and do not facilitate searching, indexing or use, internally or externally. A range of tools for document management, content management and website support do exist, but few are build with a mind to multilingual handling, archival control of content and editorial control of publishing. The system developed by Prosentient Systems for Digital Asset Management is targeted specifically to address the market gap in this area. This white paper is a conceptual overview of the software function and current literature review of the state of play in content management systems.

Further information about this software can be obtained from Edmund Balnaves, *Prosentient Systems Pty Ltd*:

Phone: 0414 737 560.

Email: ejb@prosentient.com.au

Fax: 02 9212 0899

CONTENT MANAGEMENT

A Content Management System allows documents (web pages, and potentially multiple other document formats) and metadata to be stored in a database through a simple user interface. Because the content is stored separately from the presentation design, documents can be contributed by staff who do not need to be familiar with HTML and complex publishing procedures.

Design and presentation aspects are handled by web designers who set up the appearance and functionality of the site with HTML templates. After a site has been set up and handed over it can be updated simply and consistently by pasting text into entry forms and single button clicks to preview and publish. Navigational index documents are automatically rebuilt and metatags inserted to facilitate the operation of search engines.

The separation of content also facilitates content re-use, and delivery in new technological formats (eg XML).

Prosentient systems has built a system with core capabilities in the management and re-use of digital content in a multi-lingual framework. The new Digital Asset Management System will have the following features:

- a generalised import engine capable to facilitate database capture of existing site content (images and text)
- embedded XML handling and XSL publishing capability
- considerably expanded digital asset management facilities
- inherent multi-lingual capabilities
- enhanced content re-use facilities

While there are some (very expensive) tools for content management in the market, few of them take a long term view of content management in terms of a library of digital assets, to facilitate content re-use. The new version of this software will expand the metadata and thumbnail representation of multimedia content.

Variations in information design are addressed in the ability to specify different document types. This allows for the tailoring of data entry form and the rendering of documents with different XML schemas.

One-to-many and many-to-many relationships between documents can be realised through the concept of *attachments*.

Access paths to core information content can be maximised through the combined use of generalised search engines and structured index paths to information on a site or in a publication produced from the system.

DIGITAL ASSET MANAGEMENT

Systems designed purely for website content management and publishing generally do not address to key issues: archival retention of information and multi-language handling. The focus on this new generation publishing tool is to provide a framework which:

- * facilitates flexible definition of document content (all content structures are internally described in XML)
- * provides a means of content management in a multi-language environment
- * enforces the archival management of content and the editorial process of publication.

The effective management of digital assets of all types (image, audio, text) demands a system in which the definition of content is flexible. The following screen snapshot from the epublishing system gives an indication of the logical structure of the system:

InBox	Content	Presentation	Lookups	Workflow	Auto
[Documents] [Collection] [Links] [Images] [Sounds] [Videos] [Contacts] [Surveys]					
Draft Document					Create
#	Title	Section	L	Type	
Published Document					
#	Title	Section	L	Type	
11	Alumni Registration	B - Contact Us	en	Document	Preview
12	Amazing Trees	Montage Pages - Projects	en	Document	Preview
13	Amazing Trees	Montage Pages - Showcase	en	Document	Preview
14	Anglo-Australian Observatory	Observatory - gateway	en	Document	Preview

Content is organised by “content types”, and is maintained in a “draft”, “published” and “archival” form. The attributes of each “content type” can be defined flexibly using an XML content description

LITERATURE REVIEW

Both at the national and International level, multimedia content creation is considered of immense strategic importance [1,2]. The EEC Multimedia Educational Software observatory (MESO) final report sees a “missing paradigm” for multimedia content production: it belongs neither in the processes for book publishing, film production, software development nor game development. [3]. Publishers working in the Educational environment have made a strong contribution to the production of multimedia content; they have also been among the earliest to seek a means to manage content reuse, revision and regeneration [4,5].

We use this terminology to express three different concepts. Reuse is the use of a given set of material in a number of different publications, whether they be on paper, on the WWW (as HTML files) or in a CD-ROM run-time systems. Revision is the process of changing content in a given system for the purpose of publishing a new edition of the same core materials. Regeneration is the workflow processes that need to be followed to create a revised version of a publication.

The importance of effective content re-use is highlighted by the rapid obsolescence of both technology and knowledge assets, with a longevity as little as 18 months [6]. Such a rate of obsolescence demands the building of content re-use systems that allow not only the repurposing of content into new technologies, but also the effective management of resources in a broader digital asset architecture.

Indeed, the regeneration of multimedia content presents a challenge in every stage of the publication process. The once-off publication of multimedia content might, of its nature, not require an inordinate investment in the long-term management of the content itself. However, the rapid obsolescence of information, both in its design or presentation elements and the very content itself makes the regeneration of this content with revised or updated material an early exigency, particularly in the management of Website Content [7] and in the Education arena [8].

The logical conclusion is that the content creator should in the first instance collect the component elements of a multimedia publication in a database comprising the design elements, logical structure and multimedia items themselves (text, audio, visual, etc). Multimedia capable databases are gaining some prevalence in the industry.

However, such a content database presents not only a storage conundrum, but also a retrieval dilemma. The Information Retrieval of multimedia content has been most extensively explored in the Library Science literature. Early research and trials by Lancaster [9] led to the first large bibliographic databases such as MEDLARS in the early 1970's. The 1990's saw a movement from a communication model of Information retrieval to a behavioural model of information seeking [10], with detailed analysis of this process by McCreadie and Rice [11, 12] and more recently applied in the multimedia context [13]. However, the analysis of Internet-based search behaviour has tended to remain isolated from the discipline of Information Research Behaviour, with recent papers such as Wang, Hawk et al [14] focussing on a communication model of information search behaviour on the Internet. Research constructs described by Mann [15] in 1993 in his "methods of searching" model have been extended to the analysis of discipline-specific "material mastery" in the behavioural elements of digital library usage [16]. However, the complex Interaction involved in Information Retrieval is difficult to translate to an effective User Interface, and it has been argued that such models for IR are more useful as a reference guide during the prototyping design than as a concrete instrument for conceptual design. [5, 17]. Nevertheless, over large multimedia content databases, the challenge of effective indexing of content is still being explored [18, 19]. Content indexing of video and sound archives present particular problems [20]. The behavioural elements in the process of information retrieval present a basic challenge not only in information retrieval, but also in the maximisation of the re-use of content.

The Intersection of the problems of information retrieval with the management of libraries of digital assets were highlighted by Wood et al [21]. The common use of one-shot queries in web-sites and similar systems depends greatly on the specificity and relevance of metadata associated with the content. The structured collection of metadata in itself can be a time-consuming task, and is itself very subjective. Santini and Jain[22] explore the use of Interface design to maximise the contextual information implicit in the database to enhance the process of retrieval.

The multimedia publisher commonly employs a variety of different tools for content creation [4, 23, 24]. The process of creation, therefore, can involve the complex Interaction of different content creation systems. Work by Cheung & Chanson [8] shows the added complexity of reaching and identifying multimedia resources over a heterogeneous network. They propose three different models: configuration, user control and presentation. Complexities of the software management of multimedia content are explored by Vazirgiannis et al [25] and Agoulmine et al [26]. Multiple methods of publication extend the

concern of publishers into the area of product support as well as the production and distribution of their content [7].

The production process itself is also compounded by the variety of tools and processes that go toward multimedia content generation[27,28]. The Microsoft “White Paper” on “Content Management” explores a largely linear view of the content management process [7], with little exploration of the issues of content re-use. It does, however, examine the Interrelationship between content management and workflow of the publication process. Indeed, workflow systems themselves are taking on conceptual elements of document management systems (and visa-versa) [29]. The requirements of content reuse within Multimedia Content Systems are explored by Liu and Hsu [30] from the perspective of document management systems, whilst the high cost of multimedia content publication [31] makes content reuse important. Nevertheless, the integration of workflow systems will need to go beyond the domain of the publication generation and into the arena of rights administration. Similarly, the complexity of the heterogeneous tool-set involved in the content creation makes the computer automation of processes involved in the generation of content all the more important. While the concepts underlying workflow management of business flows are well defined, they tend to be poorly defined in the editorial management of content databases, although it is clear that many concepts relating to workflow management apply; in particular:

- * availability of graphical tools
- * routing capability
- * queue management of tasks
- * management of events
- * task and process management[32].

Issues of workflow control also Intersect aspects of “trouble management” described by Agoumine et al [26] in the management of Quality of Service through the use of Quality of Service Assurance Systems. An area of vital Interest is the integration of problem management and QOS management processes within the content management process itself.

The “Content Management System” oriented to the generation of material specifically for website publication has its roots in management of complex document content [Celentano, et al, 36]. Current models for Content Management (such as those offered by Microsoft [7]) focus on the primary generation and maintenance of website content, and their origins lie in simple content delivery mechanisms for web publishing [24, 37]. Content separation from design is achieved either through Template-based approaches to content publication, or more recently through use of XSL (Extensible Style Sheet) templates Interacting with XML document formats[38]. This area of industry discussion is very topical, and is best followed through current lists such as cms-list [39] and. The explosive growth of content management systems highlights rather than diminishes the potential for some theoretical convergence in the area of content re-use, as website publishers struggle with issues of currency, navigation and editorial consistency on their sites.

The Library of Congress *National Digital Library* program illustrates the degree of intersection between digital asset management and web-based content management [40]. Their efforts to construct a web-based paradigm for information management and retrieval provide good indicators for directions that Content Management Systems themselves are likely to pursue.

Finally, obviously, the issue of content reuse is active in the education community, with active efforts to build systems that maximise the reuse of course material in CDROM and tutorial courseware preparation [5,43, 44].

References

1. OECD. Content As A New Growth Industry. *European Audiovisual Conference Birmingham, 6-8 April 1998*. http://europa.eu.int/eac/papers/oecd_en.html. 1998
2. Multimedia Educational Software Observatory. <http://europa.eu.int/comm/education/meso/>
3. Multimedia Educational Software Observatory. <http://europa.eu.int/comm/education/meso/conclusions.pdf>
4. Brett, P.A. & Nash, M. Multimedia language learning courseware: a design solution to the production of a series of CD-ROMs. *Computers & Education* Volume 32, Issue 1, January 1999, pp.19-33
5. Colbert et al. The application of process models of information seeking during conceptual design: the case of an intranet resource for the re-use of multimedia training material in the motor industry. *Proceedings of the conference on Designing Interactive systems: processes, practices, methods, and techniques*, 1997, Pages 73 - 81
6. Multimedia Educational Software Observatory. <http://europa.eu.int/comm/education/meso/forecast.pdf>
7. Reynolds, J and Kaur, Arminder. Content Management. <http://www.Microsoft.com.technet/ecommerce/contmgmt.asp>, Microsoft, 1999
8. Cheung, - S. C. and Chanson, S. T. A model-based authorware for the construction of distributed multimedia systems, *Information and Software Technology*, Volume 41, Issues 11-12, 15 September 1999, Pages 715-727
9. Lancaster, F.W. Interaction between requesters and a large mechanized retrieval system.' *Information Storage and Retrieval*, 4(2) 1968, 239-252
10. Kuhlthau, C.C. Seeking a Meaning: A process Approach to Library and Information Services. Norwood, New Jersey, Ablex, 1993.
11. Maureen McCreadie and Ronald E. Rice. Trends in analyzing access to information. Part II. Unique and integrating conceptualizations, *Information Processing & Management*, Volume 35, Issue 1, January 1999, pp 77-99
12. Kuhlthau, Carol Collier. Role of experience in the information search process of an early career information worker: Perceptions of uncertainty, complexity, construction, and sources, *Journal of the American Society for Information Science*, Volume 50, Issue 5, 1999, pp 399- 412
13. Chang et al. Multimedia access and retrieval (panel session): the state of the art and future directions. *Proceedings of the seventh ACM International conference on Multimedia* October 30 - November 5, 1999, Orlando, FL USA, pp 443 - 445
14. Wang, Peiling; Hawk, William B. and Tenopir, Carol. Users' Interaction with World Wide Web resources: an exploratory study using a holistic approach, *Information Processing & Management*, Volume 36, Issue 2, 1 March 2000, Pages 229-251
15. Mann, Thomas. *Library Research Models: a guide to classification, cataloguing and computers*. NY, OUP, 1993
16. Covi, Lisa M. Material mastery: situating digital library use in University research practices, *Information Processing & Management*, Volume 35, Issue 3, May 1999, Pages 293-316
17. Sonnenwald, Diane H. and Pierce, Linda G. Information behavior in dynamic group work contexts: Interwoven situational awareness, dense social networks and contested collaboration in command and control, *Information Processing & Management*, Volume 36, Issue 3, 1 May 2000, Pages 461-479
- 19 - Amato, Giuseppe, Rabitti, Fausto and Pasquale Savino Multimedia document search on the Web, *Computer Networks and ISDN Systems*, Volume 30, Issues 1-7, April 1998, Pages 604- 606
18. Multimedia document search on the Web, *Computer Networks and ISDN Systems*, Volume 30, Issues 1-7, April 1998, Pages 604-606
Giuseppe Amato, Fausto Rabitti and Pasquale Savino
20. Dagtas, Serhan et al. Models for motion-based video indexing and retrieval, *IEEE Transactions on Image Processing*, Volume 9, Issue 1, January 2000, Pages 88-101
21. Wood, M E J, Campbell, N.W. and Thomas, B.T. Iterative Refinement of Relevance Feedback in Content-based Digital Image Retrieval. *Proceedings, ACM Multimedia 98, Bristol, England, Sept. 12-16 1998*. ACM, NY, 1998, pp13-18
22. Santini, S & Jain, R. Beyond Query by Example. *Proceedings, ACM Multimedia 98, Bristol, England, Sept. 12-16 1998*. ACM, NY, 1998, pp.345-350.

23. Michalas, V et al. A comparison of multimedia application development platforms towards the object web, *Computer Standards & Interfaces*, Volume 22, Issue 1, March 2000, Pages 13- 26
24. Fraternali, Pierre. Tools and approaches for developing data-intensive Web applications: a survey; *ACM Comput. Surv.* 31, 3 (Sept. 1999), Pages 227 - 263
25. Vazirgiannis, et al. Specifying and authoring multimedia scenarios, *IEEE Multimedia*, Volume 6, Issue 3, 1999, Pages 24-37
26. Agoulmine et al. Trouble management for multimedia services in multi-provider environments, *Journal of Network and Systems Management*, Volume 8, Issue 1, 2000, Pages 99-123
27. Duffy, K., Content management comes of age: streamlining the information creation and delivery process, *Information Management & Technology*, Volume 33, Issue 4, 2000, Pages 183-191
28. Liu, Peiya and Hsu, Liang H. Multimedia Document Systems From Authoring perspective. *Internet and Multimedia: Systems and Applications*. Furht, B (ed). IEEE Press, 1999, pp 635- 667.
29. Collins, Philippa. Document and workflow management '99. *Management Services*. Vol 44, Issue 1, 1999, pp. 23-26
30. Hysell, Debbie. Meeting the needs (and preferences) of a diverse World Wide Web audience; *Proceedings on the sixteenth annual International conference on Computer documentation*, ACM, 1998, pp 164 - 172
31. Kelly, Brian. Will it all fit together? The need for standards and the technical challenges. *Multimedia Information and Technology*, 2000 (Vol. 26, No. 1).
32. Rojas. M Perez, T. Evaluation of Workflow-type software products: A case study. *Information and Software Technology*. Volume 42, Issue 7, May 1, 2000, pp 489-503
33. Crispin Jewitt. Digital chaos and professional standards. *Multimedia Information and Technology*, 2000 (Vol. 26, No. 2).
34. Koch, E, Rindfrey, J and Zhao, J. Copyright Protection for Multimedia Data. *Digital Multimedia and Electronic Publishing*, edited by Earnshaw, R, Vince, J and Jones, H. NY, Academic Press, 1996, pp 203-214.
35. - Cox, John E. The Changing Economic Model of Scholarly Publishing: Uncertainty, Complexity, and Multimedia Serials, *Library Acquisitions: Practice & Theory*, Volume 22, Issue 2, Summer 1998, Pages 161-166
36. Celentano, Augusto, Pozzi, Silvano and Toppeta, Donato. A multiple presentation document management system; *Proceedings of the 10th annual International conference on Systems documentation*, 1992, Pages 63 - 71
37. Yeh, Jian-Hua; Chang, Jia-Yang; Oyang, Yen-Jen. Content and knowledge management in a digital library and museum, *Journal of the American Society for Information Science*, Volume 51, Issue 4, 2000, Pages 371-379
38. W3C. Extensible Markup Language (XML) 1.0. <http://www.w3.org/TR/REC-xml>
39. Content Management Systems. <http://www.camworld.com/cms> [online discussion list]
40. Marchionini, Gary, Plaisant, Catherine and Komlodi, Anita Interfaces and tools for the library of congress national digital library program, *Information Processing & Management*, Volume 34, Issue 5, September 1998, Pages 535-555
41. Cline, Lynn. Buying electronic: the development of the electronic book market in academic libraries, *Library Collections, Acquisitions, and Technical Services*, Volume 24, Issue 2, Summer 2000, pp 312-315
42. Yeh, Jian-Hua; Chang, Jia-Yang; Oyang, Yen-Jen. Content and knowledge management in a digital library and museum, *Journal of the American Society for Information Science*, Volume 51, Issue 4, 2000, pp 371-379
- Candler C.S.; Andrews M.D. Avoiding the great train wreck: Standardizing the architecture for online curricula, *Academic Medicine*, Volume 74, Issue 10, 1999, Pages 1091- 1095