# Managing Multimedia Content Databases: toward a Model for Content Management.

**Edmund Balnaves and Jon Patrick**
Basser Department of Computer Science, University of Sydney

**ABSTRACT.** The Internet provides an effective means of dissemination of information in the Humanities. In many cases the Internet is becoming the primary or even only form for dissemination of information. In this context, the effective management of published resources becomes essential. Management of multimedia content on the Internet must deal with not only the original Content publishing process but also issues of technological obsolescence, effective reuse of the digital assets, and version control of information. Systems must address the established disciplines of effective description, classification and preservation to be more than just transient sources of information.

Many website publishing systems address only the editorial workflow control of publication and separation of content from presentation, oriented to a single runtime engine. This paper presents a model for Content Management System design that incorporates the effective identification, dissemination and re-use content in a heterogonous authoring and runtime environment. Issues of archival management and mitigation of technological obsolescence are discussed.

## INTRODUCTION

### *Content Management Systems*

The management of complex Internet sites has seen the evolution of a class of tools that aid in the site maintenance and publishing activity. These "Content Management Systems" have proliferated in the last few years, reflecting the heavy demands for editorial management of a website. As with many systems that have grown out of an immediate need within one industry segment, they have often lacked the maturity of software that has seen several years of industry development. Nevertheless, if properly conceived, this class of software offers some new opportunities for the long-term effective management and reuse of multimedia content. This paper proposes a theoretical model that should act as a standard for continued evolution of this class of software, and as a framework for evaluation of existing software offerings.

The term "content management" is a succinct description of the demands of managing multimedia resources. The term "content" is sufficiently general to describe any digital form, media or object ("content" in the sense of a descriptive label of items that are "contained within"). It also implies the elements of meta-information storage (in the sense of the "table of contents"). The term "management" can be extended to imply not only the operational management of the publishing exercise, but also the long term organisational management of the content store, the effective use of content through regeneration, repurposing and syndication, and the long term archival administration of content.

Content Management Systems as they have evolved in the Internet environment have sought to meet the immediate requirements of website publishing (Celentano 1992) in order to facilitate:

- currency of information; that is, the need for a framework through which website content can be easily updated;
- technical elements of publishing; including the mechanics of delivering changes to the website (FTP, WebDAV, dynamic delivery, etc);
- design standards; the editorial control of information content and presentation, and the management of "corporate" design standards.

Websites typically go through several generations of design. The underlying content, however, may be similar between these generations of graphical design. The transition of stable content through intergeneration graphical redesign of websites simply emphasises a trend already implicit in the early needs of Content Management Systems: the separation of content from style and presentation. Content re-use or re-purposing has a role both in the regular makeover of websites and in the syndication of content to a variety of different destinations and media. Evident here is a tension between the separation of content and style for purposes of effective management and content re-use, and the cohesion of content and style to enhance effective editing of the content. The paradigm of WYSIWYG (What You See Is What You Get) is highly intuitive for the content creator, but stands opposed to the structural need to decompose content elements and the stylistic elements for purposes of long-term management.

More recently the abstraction of site structural information as an integral element of site control has been explored (Fraternali and Paolo. 2000). A content management model, to achieve the goal of both content re-use and content preservation, will necessarily provide visual presentation of the structure of information represented in a site.

Eary software implementations of Content Management Systems focussed on the primary generation and maintenance of website content, and their origins lie in simple content delivery mechanisms for web publishing (Fraternali 1999; Yeh 2000). 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 (W3C ). This area of industry discussion is very topical, and is best followed through current lists such as cms-list (Camworld). 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.

### *Obsolescence And Technology Architecture*

The widespread availability of multimedia content authoring systems allow the content creator to draw on unprecedented resources in the process of their analytical and creative endeavour. Yet it could be argued that never has so much information

been managed so poorly by so many. The rapidity of the evolution of new means of content creation has resulted in considerable software and hardware speciation. Minor incompatibilities between versional releases of software compound themselves over time. It is sobering to reflect on the most common word processing formats just a decade ago: the ubiquitous World Perfect, Word for DOS and others. Just fifteen years earlier sees the supremacy of the Wang Word Processing Systems and the central typing pool. Radical changes in hardware architectures, operating system platforms, content storage formats are evident. This is classic entropy as it affects our information resources.

There is a fundamental architectural dissonance between the creative community and the technical community who are responsible for developing Content Management Systems. Clearly there are at least some in the arts and humanities who are trying to reach an audience through time as well as space, and who want their artistic and research creations to last through centuries rather than decades or months. Clearly, the practical life of software innovations and digital technology spans a decade at most, commonly several years only, and at the worst barely a few months.

Nevertheless, there is an inevitable move toward using new media technologies for delivery of digital content in many different forms, if only because the audience the creators and innovators are trying to reach are those using that very media.

Yet while the archival management of the book and other printed media is well understood, similar archival strategies in the new media are in their infancy. There are daunting challenges facing those who attempt to preserve in the long term digital creations of contemporary times. There is an expectation of those expending creative effort in the new media that their efforts will be available in the long term. The cost burden of long term persistence of content in current formats has traditionally been borne by Libraries. However, few libraries have the resources to sustain the research and development required to address archival issues of Information Systems. The move to dynamic and personalised content delivery places even greater emphasis on the archival management of the original content stores and their archival capability. There is a strong trend in the discussion of scholarly archiving of digital published material to place the archiving role back on the original publisher (Ekman 2000).

In the long term, addressing the issue of obsolescence requires continuous transformation of multimedia content to ensure it is available through current delivery vehicles. A central role for Content Management Systems will be the efficient management of this transformation process over time.

*Ephemera And Identification*

The Internet Universal Resource Locator (the URL) is an apt metaphor for the difficulties of unique content identification in the digital era. Digital content design and publishing systems lend themselves to a rapid rate of content design and delivery. The corollary to this is the rapidity of the entropy effect on this content. Lawrence, Pennock, Rovetz et al. (2001) examined 270,977 computer science journals, conference papers and technical reports, extracting 67,577 URL references. They demonstrated both a dramatic increase in the use of URL's in citations and the

substantial increase in broken links over time, peaking at 53% after 6 years. This again highlights the ephemeral nature of URL referencing, even within journal publications. Lawrence analyses the problem broadly as follows:

> "First, personal homepages tend to disappear when researchers move. Second, many who restructure Web sites fail to maintain old links. These problems are likely to persist without improved citation practices." (Lawrence, Pennock, Rovetz et al. 2001) p.28

Library Science has a term for loose-leaf items and irregular publications: ephemera. Ephemera often suffer from the difficulty of unique identification and the problem of description for cataloguing purposes.

The identification, access and archival characteristics of a traditional journal citation and a URL citation are very different:

| **JOURNAL CITATION** | **URL CITATION** |
|---|---|
| Content is static<br>Content can be sourced uniquely through international delivery mechanisms<br>Content is widely distributed (multiple repositories) | Content is variable<br>Content may only be available through one source (the website) |

**Figure 1: Content identification**

The URL has considerable value in the short term. There are experiments in archiving sites, such as the Pandora project and the National Library of Australia (Phillips 1998). However, this archive represents little more than 1,200 Australian sites, despite the cultural significance in the long term of preserving more of this material(Smith 1998). It could be that Internet content delivery can be regarded as essentially ephemeral. The final form of a book or a journal had the virtue of a static nature: their content is the same for all readers. Personalisation, on the other had, dispenses with any degree of finality of information delivery: the content delivery may be different for each individual. Without a fixed point of reference in which content can be thought to have reached a "final" form – that is, which is *essentially* dynamic, the issues of attempting to preserve content in its final generated form become meaningless.

The ephemeral nature of published multimedia content in its final form places greater demands on the Information Systems that were used to generate this content. As a result, the architecture of the Content Management System must incorporate strong mechanisms for unique content identification of multimedia content elements. The organisation of the content must lend itself to the effective replication of content for archival purposes and the efficient reuse of content.

*Multimedia database and Content Identification*

The multimedia publisher commonly employs a variety of different tools for content creation (Brett 1999; Michalas 2000). The process of creation, therefore, can involve the complex interaction of different content creation systems. Work by Cheung and Chanson (1999) shows the added complexity of reaching and identifying multimedia resources over heterogenous networks. They propose three different models: configuration, user control and presentation. Complexities of the software management of multimedia content are explored by (Vazirgiannis, Kostalas and Sellis 1999) and (Agoulmine 2000). 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(Reynolds 2000).

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. The regeneration of multimedia content presents a challenge in every stage of the publication process.

The Content Management Systems must therefore facilitate the effective management of the heterogenous elements of complex multimedia elements, some stored locally and some sourced externally. The multimedia content database supporting a Content Management System must support the demands of:

a) content organisation in an manner which maximises content identification and dissemination;
b) management of connectivity information for external content items and content repositories;
c) elemental organisation of content, content structure and meta-information relating to content in a manner which enhances content re-use.

*Content Indexing and Classification*

Library Science has long had effective and disciplined approaches to cataloguing and classification of information. Cataloguing has one fundamental purpose: to fully and accurately describe an item in a way that enables it subsequently to be found. The Library Catalogue describes the author, title, publisher, classification, shelf location, and sometimes the table of contents, physical book information, book identifiers and other attributes. Strict disciplines of description (*eg* Anglo-American Cataloguing Rules) allow recognition of catalogue items in a consistent manner: rules for handling multiple authors, lack of authors, consistent description of authors across multiple variants, etc.

Cataloguing therefore aims to give us a means of both describing and locating an item. Classification is an intellectual exercise of ascribing meta-information relating to an item in a structured and consistent manner. Classification schemes evolved as a response to the escalating difficulty of managing and organising large collections.

Early online indexing research and trials by Lancaster (Lancaster 1968) 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(Kuhlthau 1993), with detailed analysis of this process by McCreadie and Rice (Kuhlthau 1999; McCreadie 1999a, b)and more recently applied in the multimedia context(Collins 1999). 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 2000) focussing on a communication model of information search behaviour on the Internet. Research constructs described by (Mann 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 (Covi 1999). Over large multimedia content databases, the challenge of effective indexing of content is still being explored (Amato 1998). Content indexing of video and sound archives present particular problems (Dagtas 2000). 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.

Disciplines of cataloguing are rarely seen on the Internet. Some descriptive information that leads to the physical equivalent of the shelf location, the 'URL', is used: the most common example is the use of *meta-tags* in HTML to describe the page itself. In many ways the digital information era recalls that chaotic era prior to the adoption of universal schemes for cataloguing and classification. The "Dublin Core" was an early effort to define some standards in this area.

Text, image, audio and video content each have very different challenges in the management of large content databases. Digital content in the Internet arena has grown to some degree independent of the intellectual rigour surrounding classification systems. Meta-data classification of content needs also to address the classification of content at the elemental level, not just of content in is final form, and in a manner that enhances the distributed content identification and dissemination of the content.

The intersection of the problems of information retrieval with the management of digital collections were highlighted by Wood (1998). 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 itself can be a time-consuming task, and is itself very subjective. Santini (1998) explores the use of User Interface design to maximise the contextual information implicit in the database to enhance the process of retrieval.

Central to the role of a Content Management System, therefore, must be the flexible capture of the content itself, and the effective and consistent meta-description of the content at the elemental level. Presentation of the content structure should reflect the behavioural aspects of the content creators in information visualisation.

## CURRENT RESEARCH AND A MODEL FOR CONTENT MANAGEMENT

It is evident that the task of content management embraces a disparate range of complex functions. Yet both at the National and International level, multimedia content creation is considered of immense strategic importance (MESO; OECD. 1998). 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. (MESO). 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 and regeneration(Colbert 1997; Brett 1999). The importance of effective content re-use is highlighted by the rapid obsolescence of both technology and knowledge assets, with longevity which is as little as 18 months (MESO. 2000). Such a rate of obsolescence demands the building of content re-use systems that allow not only the repurposing of content, but also the effective conversion to new runtime environments as part of a broader digital asset architecture.

The Basser School of Computer Studies at University of Sydney is exploring a number of projects in multimedia content reuse and Knowledge Management. The following model arises from experiments in the multimedia regeneration of a multilingual thesaurus, and in the management of online educational information delivery in a multilingual context. A project in conjunction with The British Council Australia is exploring the prototype model of a Content Management System for website publishing based on theoretical requirements discussed in this paper.

Any model to fully describe a Content Management System must reflect not only the database elements of the content store but also the systems elements of the production process. As (Fraternali 1999) has indicated, Content Management in a web-site context can architecturally be described as a hybrid between hypermedia and an information system.

In particular, a systematic model for content management must address:

a) a model for content capture, separation and mark-up;
b) an effective storage model for multimedia content and for unique content identification;
c) a system for integration with authoring systems and for building associations between content stored within the Content Management System and content located externally;
d) a workflow management process for the editorial control of the publishing process;
e) an information retrieval and modelling interface to facilitate content reuse and expose the content structure;
f) a final content delivery mechanism to deploy content through heterogenous runtime environments.

Industry efforts in building Content Management Systems characteristically focus on discrete functional challenges they are addressing: original publishing, content update, efficient content delivery, etc. This paper and the research focus in this area at Sydney University is in the elaboration of a model that can form the conceptual basis for further work in this area, as well as specific issues in multimedia content generation and reuse. *Table 1* outlines this model.

**Table 1: A model for Content Management**

| |
|---|
| Layer 4: **Content Publishing Layer**<br><br>Generative systems integrated with runtime engines and presentational formats.<br>Content mapping systems (eg XAS) |
| Layer 3: **Content Information Retrieval Layer**<br><br>Search engines (eg Z39.50)<br>Clustering techniques<br>Structured Indexing techniques |
| Layer 2: **Content Storage and Identification Layer**<br><br>Unique identification and version control of content elements<br>Management of the structural elements as well as the content elements<br>Processes for Elemental decomposition of Content<br>Rule-based content-relationships<br>Structural content-relationships<br>Management of the content cross-referential structures |
| Layer 1: **Content Creation, Capture & Separation Layer**<br><br>Rule system for separation and markup of content and structure (eg XML)<br>Content capture systems and Authoring systems |

**Content Creation, Capture & Separation Layer**

Content capture is a non-trivial process in a heterogeneous media environment. The consolidation of content into a single database framework will invariably require the interaction of a range of content formats, authoring tools and content mark-up systems. With content reuse as an objective, the first prerequisite for effective content reuse is the decomposition of content into elemental content items.

In qualitative data analysis tools such as NUD*IST, the elemental decomposition of content is fundamental to the effective analysis process. In this case, the software treats this as a text unit of a line, paragraph or document. By this means, the qualitative analysis of this content (the coding) can be annotated at the relevant elemental level. Subsequent dissection and reporting of the coding hierarchy is possible. The original document is maintained intact in this framework, while a level of "meta" information is built around the relevant portions of the document. The software provides a framework focussed on interpretation of the original text through layers of coding abstraction. The software maintains the integrity of the original document and exposes for research purposes the layers of interpretation built upon the source documents.

In addition to the research narrative surrounding an original document, content publishers have a commercial requirement to track the copyright clearance levels of content they use. Similarly, from an archivist's point of view, the management of digital content presents dilemma of multiple copyright ownership (Jewitt. 2000), and the fact much relevant audio content is not in the public domain. Aspects of the social and copyright issues of multimedia publishing of academic journals explored by Cox (1998) are equally applicable to the generation and regeneration of multimedia content in an Internet framework.

The management of online educational curricula is perhaps the leading area in which the challenges of content re-use are being actively engaged. There are manifest advantages to the effective reuse of curriculum objects in common teaching areas. The cost of development of curriculum content that is designed for effective visual communication in an online environment gives a clear imperative to effective reuse of curriculum materials. The issues of architectural standardisation implied by such a goal are discussed by Colbert (Colbert 1997; Candler 1999) and many others.

SGML has played a key role in the definition of a model for content mark-up in a manner that effectively describes the separation of content, structure and specific presentation of text. The rapid uptake of XML demonstrates the desire for an effective means for the description of complex textual forms inadequately addressed by simple text stores. In literature, the Text Encoding Initiative, an SGML standard for content mark-up of Literary works serves as a framework for consistent digital encoding of large bodies of literary works. Non-textual content standards such as MusicML show the popularity of XML as a vehicle for mark-up and content identification in other media.

The rapid obsolescence of information, both in content formats and runtime engines makes the regeneration of this content with revised or updated material an early exigency, particularly in the management of Website Content (Reynolds 2000) and in the Education arena (Cheung and Chanson 1999).

The abstraction of content therefore serves the additional role of attempting to retain underlying content value in a manner which:

- engages the dilemma of long-term archival retention of the content
- provides a framework for re-generation of the same content in new or other media delivery architectures.

**Content Storage and Identification Layer**

Nevertheless, content mark-up goes only partway toward the construction of systems for subsequent content re-use. Architectural considerations are of profound importance in the management of content. While database vendors are increasingly providing database tools which integrate the management of multimedia audio and video content, the efficiencies of the traditional OS file system often make it the popular choice for storage of audio, video and image content itself, while meta-data referencing the content is retained in database tables.

In this context, SGML as a standard is only part of an overall solution. A long-term tension between the communities building SGML-like architectures such as XML, and the database communities supporting large, high performance databases, has not lent to a happy resolution of problems in the area of very large databases of digital content. If anything, the problems of obsolescence are exacerbated rather than aided by current trends.

The hierarchical network model for content mark-up represented by SMGL needs to be mapped to a database framework which:

a) Uniquely identifies elemental components of the content for purposes of syndication and re-use;
b) For purposes of effective editorial control tracks the relationships between content as it has been used in various different forms.

A key area for further research is the effective database integration of heterogenous multimedia sources in a manner which not only addresses the storage issues of that content but the consistent content identification issues. To address the problem of the ephemeral nature of digital content publishing in its final form, content databases need to address issues of unique content identification in a manner that facilitates content exchange between various content stores. This is a key prerequisite to realising archival value of content in the long term. This is an issue of standards not only in the content storage but also at the national level for the consistent meta-identification of the content in a way that allows archival content repositories to be established. The Book Deposit scheme at National Libraries has been an effective mechanism for ensuring the preservation of key published content in print form within the national collection. The complexities of content identification and the technological incompatibilities at the content capture layer make it all the more important for the development of meta-identification standards that go beyond the limited Dublin Core.

**Content Information Retrieval Layer**

Information retrieval architectures such as Z39.50 provide a framework for the distributed interrogation of information resources, and this standard has recently been extended to incorporate the “digital library object”.

While distributed models for document interchange exists, they have yet to address complex issues of multilingual content management in the context of the heterogeneity of multimedia forms. In the management of multimedia resources, the integral relationship between the technological platform for the information delivery and the digital content present considerable problems.

The information retrieval layer of a Content Management System must realise several goals:

a) the effective discovery of content that may be distributed across different content stores, in different languages and storage architectures

b) the modelling of content toward effective content re-use,

c) the cost-economical persistence of access to content.

While standards such as Z39.50 have yielded good results in the content dissemination and information retrieval of academic collections they require considerable development if they are to address the ephemeral nature of content in the multimedia environment.

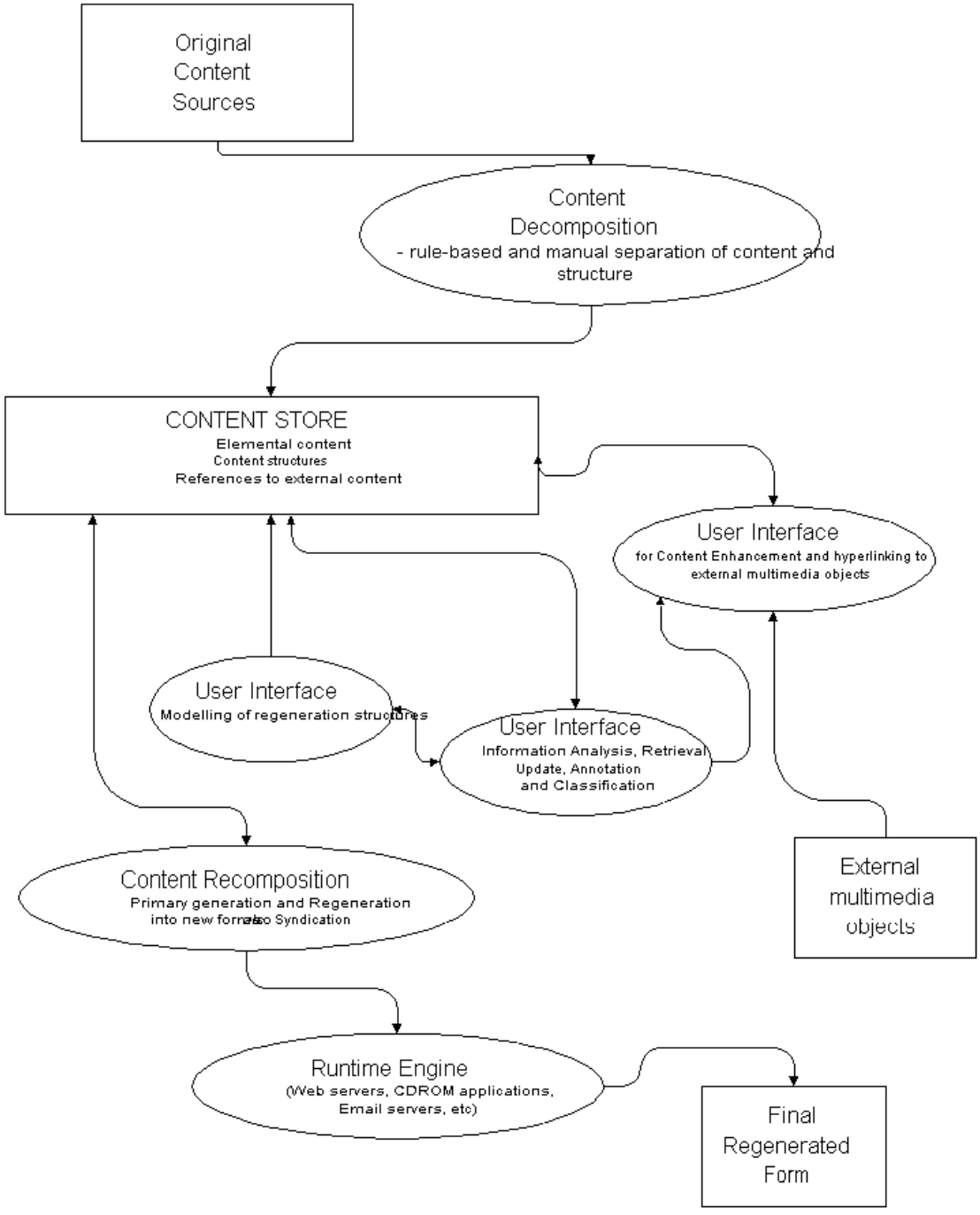


**Figure 2: The process flow for content re-use**

**Content Publishing and Reuse**

The diversity of authoring tools that feed into the multimedia production process compounds the difficulty of content reuse (Agoulmine 2000; Duffy 2000). Industry models such as the Microsoft "White Paper" on "Content Management" explore a largely linear view of the content management process (Reynolds 2000), with little exploration of the issues of content re-use. The requirements of content reuse within Multimedia Content Systems are explored by Liu (1999) from the perspective of document management systems, whilst the high cost of multimedia content publication(Kelly 2000) make content management vital. 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 heterogonous 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(Rojas. M Perez 2000).

Integral to the production process in the Content Management System is the workflow control of the editorial process. This can involve the separation of the roles of the designer, the content composer and the editorial supervisor. Such editorial management enhances the distribution of content capture throughout an organisation while maintaining the content and design standards essential for publication.

Online Educational delivery continues to be a source of considerable experimentation in the construction of educational objects that can easily be deployed in a variety of educational delivery environments, and to a range of target audiences. Just as personalisation within a commercial web delivery framework has been a focus of website design, so we are also likely to see greater efforts at the personalisation of higher educational delivery mechanisms. Universities with a focus on building campus-wide models for effective capture of educational objects will be the best placed to begin the delivery of effective, personalised online educational delivery.

Website content delivery platforms can broadly be characterised in two models of operation:

a) a generative process that produces a "static" result on the media
b) a dynamic generative process that delivers the final presentation of the content on demand.

Performance and architectural issues are the driving factors behind these two models for content delivery. A traditional website content serving static pages or with minimal dynamic generation of content still represents the most speed-efficient runtime engine (Mendes and Almeida 1998). The web server, acting essentially as a content file server, needs little interpretive intervention to deliver the content. Some web content delivery platforms, such as Vignette, have made a virtue of proprietary caching and generative mechanisms that deliver content as quickly as possible within the constraints of dynamic regeneration of the content in its final form. Dynamic online generation of content has considerable advantages for currency of content and personalisation, but requires itself a new layer of architectural support to meet the higher server load demand (Anderson 2001).

Whether dealing with a static or dynamic model of content delivery, the underlying content management issues are similar – the difference between dynamic and static content delivery is essentially is essentially one of the timing for content recomposition. The final, delivered, form of the content may be highly proprietary in form and transient in its expected longevity. The Content Management System must support both generation of content in its original form and translation to new runtime environments. An extension of this regenerative process is the long-term persistence of content in current delivery formats to address the issue of technological obsolescence.

*Figure 2* describes the process flow for the content regeneration cycle, through the stages of original content capture, of content enhancement through inter-linking with external resources, the stages of content identification, information modelling and retrieval, and finally the regenerative process for content re-use in various runtime contexts. This process flow has been tested in the context of a Content Management System used for Website publishing and in a project for regeneration of a multimedia language thesaurus.

## CONCLUSION

Management of complex and dynamic content websites has seen the evolution of a class of software to meet the management dilemma for effective control of these websites: the “Content Management Systems”. It is the proposition of this paper that a well-described theoretical model for this class of software will enrich the potential for these systems to meet increasing demands for content reuse, archiving and version control. To achieve this, a cross-disciplinary model bringing in concepts of database management, information retrieval and some elements of workflow control is proposed. Such a model can serve as a basis to guide further development of systems in this area and as a benchmark for judging the functional completeness of products.

**REFERENCES**


AGOULMINE, N., DRAGAN D., GRINGEL T, HALL, J. ROSA E. AND TSCHICHHOLZ M. (2000): Trouble management for multimedia services in multi-provider environments. *Journal of Network and Systems Management* **8**(1):99-123.

AMATO, G., RABITTI, FAUSTO AND PASQUALE SAVINO (1998): Multimedia document search on the Web. *Computer Networks and ISDN Systems* **30**(1-7).

ANDERSON, K.M. (2001): Data scalability in open hypermedia systems. *Proc. Proceedings of the tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots*, February 21 - 25, 1999, Darmstadt Germany, 27-36, IEEE.

BRETT, P.A.N.M. (1999): Multimedia language learning courseware: a design solution to the production of a series of CD-ROMs. *Computers & Education* **32**(1):19-33.

CAMWORLD Content Management Systems. http://www.camworld.com/cms.

CANDLER, C.S.A.M.D. (1999): Avoiding the great train wreck: Standardizing the architecture for online curricula. *Academic Medicine* **74**(10):1091-1095.

CELENTANO, A., POZZI, SILVANO AND TOPPETA, DONATO (1992): A multiple presentation document management system. *Proc. Proceedings of the 10th annual international conference on Systems documentation*, 63 - 71.

CHEUNG, S.C. and CHANSON, S. (1999): A model-based authorware for the construction of distributed multimedia systems. *Information and Software Technology* **41**(11-12):715-727.

COLBERT, M., PELTASON, CHRISTOF ,FRICKE, ROLF AND SANDERSON, MARIANA (1997): 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. *Proc. Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques*,, 73-81.

COLLINS, P. (1999): Document and workflow management '99. *Management Services* **44**(1):23-26.

COVI, L.M. (1999): Material mastery: situating digital library use in University research practices. *Information Processing & Management* **35**(3):293-316.

COX, J.E. (1998): The Changing Economic Model of Scholarly Publishing: Uncertainty, Complexity, and Multimedia Serials. *Library Acquisitions: Practice & Theory* **22**(2):161-166.

DAGTAS, S.A.-K., WASFI; GHAFOOR, ARIF; KASHYAP, RANGASAMI L. (2000): Models for motion-based video indexing and retrieval. *IEEE Transactions on Image Processing* **9**(1):88-101.

DUFFY, K. (2000): Content management comes of age: streamlining the information creation and delivery process. *Information Management & Technology* **33**(4):183-91.

EKMAN, R.H. (2000): Can Libraries of Digital Materials Last Forever? *Change* **32**(2):23-35.

FRATERNALI, P. (1999): Tools and approaches for developing data-intensive Web applications: a survey. *ACM Computer Survey* **31**(3):227 - 263.

FRATERNALI, P. and PAOLO., P. (2000): Model-Driven Development of Web Applications: The Autoweb System. *ACM Transactions on Information Systems* **18**(4):323-341.

JEWITT., C. (2000): Digital chaos and professional standards. *Multimedia Information and Technology* **26**(2):148 - 150.

KELLY, B. (2000): Will it all fit together? The need for standards and the technical challenges. *Multimedia Information and Technology* **26**(1):67-72.

KOCH, E., RINDFREY, J AND ZHAO, J. (1996): Copyright Protection for Multimedia Data. In *Digital Multimedia and Electronic Publishing*. 203-214. EARNSHAW, R., VINCE, J AND JONES (ed) NY, Academic Press.

KUHLTHAU, C.C. (1993): *Seeking a Meaning: A process Approach to Library and Information Services*. NJ, Norwood.

KUHLTHAU, C.C. (1999): 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* **50**(5):399-412.

LANCASTER, F.W. (1968): Interaction between requesters and a large mechanized retrieval system. *Information Storage and Retrieval* **4**(2):239-252.

LAWRENCE, S., *et al.* (2001): Persistence of Web References In Scientific Research. *Computer* **34**(2):26-31.

LIU, P.A.H., LIANG H (1999): Multimedia Document Systems From Authoring perspective. In *Internet and Multimedia: Systems and Applications*. FURHT, B. (ed), IEEE Press.

MANN, T. (1993): *Library Research Models: a guide to classification, cataloguing and computers*. NY, OUP.

MCCREADIE, M.A.R., RONALD E. (1999a): Trends in analyzing access to information. Part I: cross-disciplinary conceptualizations of access. *Information Processing & Management* **35**(1):45-76.

MCCREADIE, M.A.R., RONALD E. (1999b): Trends in analyzing access to information. Part II: Unique and integrating conceptualizations. *Information Processing & Management* **35**(1):77-99.

MENDES, M.A.S. and ALMEIDA, V.A.F. (1998): Analyzing the impact of dynamic pages on the performance of Web servers. *Proc. CMG Proceedings, Proceedings of the 1998 24th International Conference for the Resource Management & Performance Evaluation of Enterprise Computing Systems, CMG. Part 1 (of 2), Dec 6-11 1998*, Turnersville, NJ, USA, 539-547, CMG.

MESO. Multimedia Educational Software Observatory. http://europa.eu.int/comm/education/meso/.

MICHALAS, V.E.A. (2000): A comparison of multimedia application development platforms towards the object web. *Computer Standards & Interfaces* **22**(1):13-26.

OECD. (1998): Content As A New Growth Industry. *Proc. European Audiovisual Conference 6-8 April 1998.*, Birmingham.

PHILLIPS, M.E. (1998): Tomorrow's incunabula: preservation of Internet publications. *Lasie* **29**:5-10.

REYNOLDS, J.A.K., ARMINDER (2000): Content Management.

ROJAS. M PEREZ, T. (2000): Evaluation of Workflow-type software products: A case study. *Information and Software Technology* **42**(7):489-503.

SANTINI, S.J., R. (1998): Beyond Query by Example. *Proc. Proceedings, ACM Multimedia 98, Bristol, England, Sept 12-16 1998*, Bristol, 345-350., ACM.

SMITH, W. (1998): Lost in cyberspace: preservation challenges of Australian Internet resources. *Lasie* **29**:6-25.

VAZIRGIANNIS, M., KOSTALAS, I. and SELLIS, T. (1999): Specifying and authoring multimedia scenarios. *IEEE Multimedia* **6**(3):24-37.

W3C Extensible Markup Language (XML) 1.0. http://www.w3.org/TR/REC-xml.

WANG, P.H., WILLIAM B. AND TENOPIR, CAROL (2000): Users' interaction with World Wide Web resources: an exploratory study using a holistic approach. *Information Processing & Management* **36**(2):229-251.

WOOD, M.E.J., CAMPBELL, N.W. AND THOMAS, B.T (1998): Iterative Refinement of Relevance Feedback in Content-based Digital Image Retrieval. *Proc. Proceedings, ACM Multimedia 98, Bristol, England, Sept 12-16 1998*, Bristol, 13-18, ACM.

YEH, J.-H.C., JIA-YANG; OYANG, YEN-JEN (2000): Content and knowledge management in a digital library and museum,. *Journal of the American Society for Information Science* **51**(4):371-379.