CERTAIN INVESTIGATIONS ON BUILDING E-CONTENTS FOR DIGITAL LIBRARY DATABASE ENVIRONMENT: AN EXPERIMENTAL APPROACH

M.PALANISAMY¹, Dr.N.TAMILSELVAN², M.SIVAMANI³

¹Chief Librarian, SNS College of Technology, Coimbatore, Tamilnadu, India
²Chief Librarian, Rathinam Institute of Technology, Coimbatore, Tamilnadu, India.
³Librarian, P.K.R Arts College for Women, Gobichettipalayam, Tamilnadu, India.

ABSTRACT

Internet and a broad spectrum of web based applications have added a new dimension to the technologies and techniques employed for Content Management System (CMS). E-Content refers to the creation, delivery, and management of training in electronic environment. The convergence of e-content and digital libraries creates challenge to be solved not only at technical level but also at the institutional level. The handling of e-content management systems in conjunction with other prominent web applications and the perfectly smooth integration of content creation, search, access, updating etc., in terms of the different technologies is still inconclusive and remains as an open ended issue. This paper discusses about the trends of CMS and evolution (a tool based database oriented approach) of such a unified CMS for e-content in the context of digital libraries.

KEYWORDS:
Digital library, Content management system, CADDTIE, E-content.

Content Management Systems (CMS) will be a major vehicle for this digital library automation process[1]. Generally, the term "Content Management System" is synonymous with "Web Content Management Systems", further confusing the consumer.

There are no universally accepted standards or features constituting what CMS systems should contain or do. Thus, the definition of a CMS has become quite blurred when comparing various vendor solutions. Furthermore, the boundaries of the CMS system overlap considerably with document management systems, knowledge management systems, enterprise application integration systems, e-commerce systems and portals.

Additionally, there is a growing recognition that these systems also share many common features with groupware products and virtual learning environments.

1.1 Definition of Content

Content is in essence any type or unit of digital information that is used to populate a page - web page or otherwise it can be text, images, graphics, video, sound etc. In other words, anything is likely to be published across the Internet, intranet and/or extranet.
1.2 Objective

During collection, information is either created or acquired. It is then converted into a master format (such as XML) and segmented into discrete chunks called content components. Components are metadata containers that make it easier to organize, store and retrieve the content.

Content is managed within a repository that consists of database records and files containing content components plus administrative data (i.e. the system’s users). To make content available, the content management system publishes to target publications such as websites, printable documents and email newsletters.

2. PROPOSED MODEL

The proposed CADDTIE model is the generic process used by instructional designers and digital library content developers. The seven phases-Create; Analysis, Design, Development, Testing, Implementation and Evaluation represent a dynamic, flexible guideline for building effective digital library content tools.

Most of the current instructional design models are spin-offs or variations of the CADDTIE model; other models include the Dick & Carey and the Kemp ISD models. One commonly accepted improvement to this model is the use of rapid prototyping.

These ideas receive continual or formative feed back while instructional material is being created. This model attempts to save time. For example, the CADDTIE model was used in the framework for helping create new research topics in learning technology in digital library users.

2.1 Methodology

The New CADDTIE model consists of seven stages- Create, Analysis, Design, Development, Testing, Implementation and Evaluation. The Fig.1 exposes the New CADDTIE model.

The information gain ratio is just the ratio between the information gain and the intrinsic value:

\[
IGR (E_x, a) = \frac{IG}{IV}
\]

Evaluation E-content formula is Gain Ratio (GR) =

\[
\frac{(Post-test score (M) - pre-test score (n))}{(max possible score (N) - pre-test score (n))} \times 100
\]

GR= \((M-n)/(N-n)) \times 100

1. Ability to process the information and to report.
2. Ability to critically evaluate the information and make comments.
3. Ability to solve the problem.
4. Ability to apply the knowledge to real life situation.
5. Ability to relate the subject/problem in large and global content so as to see full the implications of the issue
6. Ability to develop a world view which helps the development of mankind.

2.2 Content search

Content Management contributes to the effective management of various kinds
of content by combining rules, process and workflows in such a way that centralized webmasters and decentralized web authors/editors can create, edit, manage and publish all the content of a digital library web page in accordance with a given framework or requirements.

To make content available, the content management system publishes to targeted publications such as websites, printable documents and email newsletters.

During collection, information is either created or acquired. It is then converted into a master format (such as XML) and segmented into discrete chunks called content components as shown in Fig.2. Components are metadata containers that make it easier to organize, store and retrieve the content.

3. CMS in CMA & CDA

A Content Management System (CMS) is a system used to manage content, typically for a website. Content management systems are sometime referred to as Web Content Management Systems (WCMS). Generally, a CMS consists of two elements: the Content Management Application (CMA) and the Content Delivery Application (CDA) as shown in Fig.3. The CMA element allows the content manager or author (who may not know Hypertext Markup Language-HTML) to manage the creation, modification and removal of content from a website, without needing the expertise of a webmaster.

The CDA element uses the information, compiling it to update the website. The features of a CMS system vary, but most include web-based publishing, format management, revision control, as well as indexing, search and retrieval.

The web-based publishing feature allows individuals to use a template or templates approved by the organization, as well as wizards and other tools, to create or modify web content.

- The format management feature allows documents (including legacy electronic documents and scanned paper documents) to be formatted into HTML or Portable Document Format (PDF) for the website.
- The revision control feature allows content to be updated to a newer version or restored to a previous version. Revision control also tracks any changes made to files by individuals [3].
• A CMS system indexes all the data within an organization.

• Individuals can search for data.

• The CMS system retrieves data using keywords.

A content management system starts with a purpose and a set of target publications [4] from these, a set of content components is derived that serve the stated purpose and that can be combined to create any of the target publications. A metadata framework is built around these components to allow them to be created, managed, and drawn into publications by a staff whose actions are guided by a set of codified procedures called workflows.

Fig. 4 Content processing model

A CMS enables a variety of centralized technical and de-centralized non-technical staff to create, edit, manage and finally publish a variety of content (text, graphics, video etc), under the constraint of a centralized set of rules, processes (Fig.4) flows that ensure a coherent, validated website appearance.

A content management system is needed when there is too much information to collect, manage and publish by hand.

A CMS may also provide tools for one-to-one marketing, giving webmasters the ability to tailor their websites’ content, and advertise to a user’s specific characteristics by using information provided by the user, or gathered by the site (for example, a particular user’s page sequence pattern). For example, if you visit a search engine and search for digital camera, the advertising banners will belong to businesses that sell digital cameras.

3.1 CMS process in digital library

A CMS can be broken down into four categories by function in digital library. Content collection or authoring, workflow, storage or management, and publishing. CMS system manages the flow of content from authoring to publishing by using a plan of workflow [2] as shown in Fig. 5 and by providing content storage and integration.

3.1.1 Collection/Authoring

The collection system as shown in Fig. 6 includes the tools, procedures and staff that are employed to gather content, and provide editorial and metadata processing.

The content collection process consists of adding new components to the existing repository.
Content collection can be divided into these categories:

3.1.2 Authoring

This is the process of creating content from scratch. Authors almost always work within an editorial framework that allows them to fit their content into the structures of a target publication. Authors should also be made aware of the metadata framework that has been developed for the downstream use of the content. Authors are in the best position to tag their own creations with metadata information. So authors should be encouraged and empowered to implement the metadata framework within their content as much as possible.

3.1.3 Acquiring

Acquiring as shown in Fig.8 is the process of gathering information that wasn't originally created for your CMS. This process might be partly manual or fully automated.

3.1.4 Aggregation

This is the process of gathering pre-existing content together for inclusion in the system. Aggregation as shown in Fig.9 is generally a process of format conversion followed by intensive editorial processing and meta-tagging. The conversion changes the formatting of the content, while the editorial processing serves to segment and tag the content for inclusion in the repository. Obviously, the closer the original content conforms to the standard specified in the content management system’s framework - both its editorial structure and its metadata structure and the meta information that has been entered -, the easier the aggregation is.

3.1.5 Conversion

This is the process of changing the metadata. Structure of the content (i.e., its tagging structure). During this process, the
structural and the format-related codes must be both handled. A conversion problem may appear while identifying structural elements (sidebars or footers, for example) that have only format codes marking them in the source content as shown in Fig.10. Another problem may appear while transforming formatting elements that don’t exist in the target environment.

Fig.10 Content converting [8]

4. RESULTS AND DISCUSSIONS

The average rating of user satisfaction is 9.1 and the average rating of usefulness of the tool is 9.2. Both numbers are rated on a scale of 1 to 10 in which 1 is the lowest score.

Table 1.1 Gain ratio result

<table>
<thead>
<tr>
<th></th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion Rate</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Task Time</td>
<td>11.3</td>
<td>11.4</td>
<td>15.1</td>
</tr>
<tr>
<td>(min)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Closeness to Expertise</td>
<td>0.483</td>
<td>0.752</td>
<td>0.712</td>
</tr>
<tr>
<td>Mean Goal Achievement (%)</td>
<td>97.4</td>
<td>97.4</td>
<td>98.2</td>
</tr>
</tbody>
</table>

From the numbers shown in Table 1.1, it appears that our users are highly satisfied with the tool and consider this tool highly useful for building digital libraries based on CADDTIE model.

In addition to satisfaction rating, participants rate their understanding of CADDTIE before and after using the tool.

Mean values of their rating are shown in Fig.11 is the best rating, which means the participant understands CADDTIE very well. 1 is the worst rating, which means the participant does not understand the theory at all. Statistical 87 analysis (\( t \) test with \( \leq 0.05 \)) shows that the mean value of post-understanding is greater than that of pre-understanding.

It is observed that the tool is helpful to increase the understanding of the CADDTIE Model.

5. CONCLUSIONS

Information Technology has made access to information easier, in the sense that all digital information, such as databases, full-text journal articles etc., can be accessed through computers on the network anywhere at any time.

The concept of Electronic Library gave a new dimension to the libraries and it is the time to gradually switch to Digital library. E-Collections building is one of the initial tasks. The libraries planning for E-Collections building should have a Local Area network, Content creation section, and the necessary Software/Hardware infrastructure
REFERENCES

4. Open Source CMS
5. CMS Watch
6. www.wikipedia.org