CONTENT MANAGEMENT SYSTEM FOR ART LAB, SDSU

A Thesis
Presented to the
Faculty of
San Diego State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Computer Science

by
Chetan Gowda
Summer 2012
SAN DIEGO STATE UNIVERSITY

The Undersigned Faculty Committee Approves the

Thesis of Chetan Gowda:

Content Management System for Art Lab, SDSU

[Signatures]

Carl Eckberg, Chair
Department of Computer Science

Kris Stewart
Department of Computer Science

Leigh Colnoir
School of Art, Design and Art History

5/08/2012
Approval Date
DEDICATION

To my family & friends.
ABSTRACT OF THE THESIS

Content Management System for Art Lab, SDSU
by
Chetan Gowda
Master of Science in Computer Science
San Diego State University, 2012

A website is a set of related webpages containing content such as text, image, video, audio, etc. A website is hosted on at least one web server, is accessible via a network such as the internet or a private local area network through an internet address known as a Uniform Resource Locator. All publicly accessible websites collectively constitute the World Wide Web (web). In today’s world the web is used as the primary platform for information exchange. The web is considered as mine house of knowledge. We can literally find information on almost everything that exists.

Given the web’s increasingly important role in society, access to the web is vital for people with disabilities. In order to make a website accessible, it needs to follow some of the guidelines laid out the Americans with Disabilities Act (ADA) compliance, Section 508. The research incorporates such ADA standards into a publicly accessible content management system (CMS), so that people with access challenges can also easily access the website. This CMS provides features for uploading art related images, metadata, creation and generation of PDF format slide shows from a selected set of images. The main purpose of this research is to provide insight into how ADA compliant CMS can be built and how innovative features can be combined with usability.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2 Research Objective</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.3 Thesis Organization</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE SURVEY</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2.1 Origin and Importance of Web Accessibility</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2.2 Technological Survey</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2.3 Origin and Overview of Web Based Content Management Systems</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.3.1 Repository in File System or Database</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2.3.2 Multichannel Delivery</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2.4 Origin and Advantages of PDF</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>IMPLEMENTATION</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>3.1 Architecture</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>3.2 Model View Controller Pattern</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>3.2.1 Controller</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>3.2.2 Model and Entity Classes</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>3.3 Image Processing</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>3.4 Search Functionality</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>3.5 HTML to PDF Conversion</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>3.6 Environmental Setup for the Application</td>
<td>26</td>
</tr>
</tbody>
</table>
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1.</td>
<td>Evolution of HTML</td>
<td>13</td>
</tr>
<tr>
<td>Table 2.2.</td>
<td>History of CSS</td>
<td>14</td>
</tr>
<tr>
<td>Table 2.3.</td>
<td>CMS Explained</td>
<td>18</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

PAGE

Figure 2.1. Browser main components .................................................................6
Figure 2.2. Evolution of HTML and CSS.............................................................15
Figure 3.1. Architecture diagram of the application........................................20
Figure 3.2. MVC collaboration diagram............................................................21
Figure 3.3. Screenshot showing thumbnail views of uploaded images...............25
Figure 3.4. Screenshot showing the metadata search functionality..................26
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADA</td>
<td>Americans with Disabilities Act</td>
</tr>
<tr>
<td>CMS</td>
<td>Content Management System</td>
</tr>
<tr>
<td>CRUD</td>
<td>Create, read, update and delete</td>
</tr>
<tr>
<td>CSS</td>
<td>Cascading Style Sheet</td>
</tr>
<tr>
<td>HTML</td>
<td>Hyper Text Markup Language</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hypertext Transfer Protocol</td>
</tr>
<tr>
<td>PDF</td>
<td>Portable Document Format</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
<tr>
<td>WAI</td>
<td>Web Accessibility Initiative</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to convey my heartfelt thanks to Dr. Carl Eckberg for his continuous support and encouragement towards my thesis work. Further, I would like to thank Dr. Kris Stewart for her patience, inputs and feedback provided while performing this thesis project. Next, I would thank Leigh Cotnoir, for giving me an opportunity to work on this thesis project. The interest and technical guidance provided by Leigh was invaluable.

This is an opportunity to thank my family for showering unconditional love and care as always. I would like to thank all my friends who have helped me in this project.
CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

With the invention of the internet more and more information is consumed online. The internet has successfully replaced traditional media and has become the main media for information delivery. Information on the internet is delivered via webpages that are formatted in Hyper Text Markup Language (HTML). These webpages are accessed via clients like web browsers, which can interpret HTML and display formatted information. The most popular protocol to exchange information on the internet is Hyper Text Transport Protocol (HTTP) that is built on top of technologies like Transport Control Protocol (TCP).

Origin of the internet dates back to August 1962. J.C.R. Licklider of MIT envisioned a globally interconnected set of computers through which everyone could quickly access data and programs from any site. The internet has evolved and developed immensely since 1962. Originally its connectivity was limited to a few computers inside the university, and today through internet we can connect globally.

The World Wide Web (web) is the complete set of documents residing on all internet servers that use the HTTP protocol. These documents are accessible through a web browser and can be navigated by using hyperlinks that have been implemented using HTML. In short, the World Wide Web is a web of interconnected pages and files that we are so used to browsing today [1].

As the web becomes an increasingly important resource in many aspects of life: education, employment, government, commerce, health care, recreation, and more, it is essential that the web is accessible in order to provide equal access and equal opportunity to people with disabilities. An accessible web can also help people with disabilities more actively participate in society.

The web offers the possibility of unique access to information and interaction for many people with disabilities. That is, the accessibility barriers to print, audio, and visual media can be much more easily overcome through web technologies. The Federal
Government addresses accessibility at the workplace in both the Americans With Disabilities Act (ADA) and the Federal Rehabilitation Act. In brief, both legal documents can be interpreted to apply towards the use of technology. The government amended the Federal Rehabilitation Act in 1998. One component of this amendment is Section 508. Section 508 is the basis for the movement towards accessibility on the web for the United States government [2].

The W3C is a group dedicated to the evolution of a more usable web. The W3C was created in 1994. Web accessibility has been a major undertaking of the group. Many of the recommendations that have been made by the W3C have become the foundation for governments at the country and state level, as well as academic institutions and private industry. W3C was created to ensure compatibility and agreement among industry members in the adoption of new standards [3].

Prior to its creation, different vendors, increasing the potential for inconsistency between webpages, offered incompatible versions of HTML.

The World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI) employed these guidelines to produce HTML Author Guidelines. According to the guidelines, measures for improving accessibility fall into three categories [4]:

1. **Structure**: HTML documents should use markup to convey meaning and less for format and layout pages.
2. **Navigation**: Authors should support keyboard only navigation and methods to facilitate orientation.
3. **Alternative Content**: Authors should always provide alternative ways to access information presented with images, sounds, applets, and scripts.

Eventually the final aim of WAI is that anyone using any type of web browsing technology must be able to navigate and interact with the web and to get a full understanding of the provided information without discrimination of any sort.

Here is the different kind of people who may access the web:

- Visual Disabilities: Blindness, color Blindness, other visual impairments
- Physical Disabilities: Weakness, limitation of muscular content
- Hearing Disabilities: deafness, people who had hearing problems
- Cognitive Disabilities: dyslexia, attention deficit disorder
So while building a web application, all categories of end-users has to be taken consideration.

1.2 Research Objective

The goal is to develop a content management system (CMS), which contains features such as image upload, image retrieval, edit and resize options, PDF slide creation, integrating these features with web accessibility standards. Even though there are such web applications available in the market, most of them are paid and also most of them lack ADA compliance. The intent here is to develop an ADA complaint CMS for the Art-Lab, SDSU. As the term suggests CMS allows the users with appropriate credentials to upload and manage content. Content in this scenario refers to images and metadata associated with them. There were certain challenges which required more effort and focus, mainly tasks such as conversion of images embedded in HTML to PDF format.

Some of the main goals of the application are below:

- To develop a web application that allows faculty members to upload and edit images and metadata, from any where in the world via internet.
- All the images uploaded are by default resized as thumbnails, medium and large.
- People with disabilities can access and decipher our web application.
- This application allows the students to access the image and metadata content and download the slides in PDF format.
- Consists of an admin panel, which allows Faculty/Admin to perform CRUD (create, read, update and delete) operations.
- It also incorporates a metadata search tab, which allows the users to search images by typing the key words, i.e. metadata info. Multiple criteria can be given in order to retrieve an image.

1.3 Thesis Organization

This thesis, deals with the CMS for Art Lab, SDSU. It begins with introducing the need and growth for internet over the years. Chapter 2 discusses about web accessibility background and its importance. Also it reviews the different internet technology used in this application. It sheds light on nature and origin of CMS. Chapter 3 explains the architecture and features of the application. The details of the application will be discussed in detail there. Further, Chapter 4 is about different testing scenarios followed to ensure that the application is working as expected. Chapter 5 is about Results and obstacles faced during the
development. Chapter 6 talks about the future enhancements, which can be incorporated. Finally, Chapter 7 concludes by giving the importance of web accessibility and usability features.
CHAPTER 2

LITERATURE SURVEY

Life today is highly dependent on computers; they perform important work and are found everywhere, from homes, to police stations, government institutions and military facilities. However, the computer would have never been so popular and widely spread if it was not for the internet.

The explosive growth of the internet over the last decade is attributed to two basic reasons - the non-centralized management of its development and the non-proprietary nature of its main functional units - the internet protocols. Since most of the content available in the internet are Open Source, everyone using the internet has an opportunity to contribute and implement their ideas.

The research and education community continues to develop and use advanced internet features. The internet continues to grow, driven by ever-greater amounts of online information and knowledge, commerce, entertainment and social networking. Over its history, the internet has functioned as mega storehouse of information. This feature can be credited to massive storage space available across all the web servers and the protocol like Transmission Control Protocol and Internet Protocol (TCP/IP) used to connect to different systems to access this information across from various geographic locations. The term internet means its interconnection of different computer networks. The web can be referred to as gigantic internet tool or service. Tim Berners-Lee can be credited as the one of creators of the web in 1991. Since the web is based on HTML users are allowed access information in Multimedia format.

Simplest unit is the webpage, which is primarily a document encoded in HTML format that can be accessed by using any web browser.

Web browsers are probably the most widely used software. The browser main functionality is to present the web resource that was requested by the user. And it does that by internally requesting it from the server and displaying it on the browser window. The resource is usually an HTML document, but may also be a PDF, image, or other file type.
The location of the resource is specified by the user using an Uniform Resource Identifier (URI).

The way the browser interprets and displays HTML files is specified in the HTML and CSS specifications. These specifications are maintained by the W3C organization, which is the standards organization for the web. It ensures that all the webpages are built according to web standards and they remain accessible. See Figure 2.1 for different components present in a browser [5].


Each of the components in a browser plays a specific role. An explanation of the roles of different components are as follows:

- **The user interface** - this includes the address bar, back/forward button, bookmarking menu. This includes all the components, which we see on a web browser except the main central area, which is used to display the requested content.

- **The browser engine** - marshals the actions between the UI and the rendering engine.
The rendering engine - responsible for displaying the requested content. For example if the requested content is HTML, it is responsible for parsing the HTML and CSS and displaying the parsed content on the screen.

Networking - used for network calls, like HTTP requests. It has platform independent interface and underneath implementations for each platform.

UI backend - used for drawing basic widgets like combo boxes and windows. It exposes a generic interface that is not platform specific. Underneath it uses the operating system user interface methods.

JavaScript interpreter - used to parse and execute the JavaScript code.

Data storage - this is a persistence layer. The browser needs to save all sorts of data on the hard disk, for examples, cookies.

2.1 ORIGIN AND IMPORTANCE OF WEB ACCESSIBILITY

An information technology system is accessible to people with disabilities if it can be used in a variety of ways that do not depend on a single sense or ability. Here is the brief information regarding how ADA compliance originated. Section 504 of the U.S. Rehabilitation Act of 1973 states that [6]:

- Establishes that disability rights are a form of civil rights and therefore covered by the 14th Amendment of the U.S. Constitution.
- Mandates that institutions receiving federal funds provide equal access to their programs.
- Uses total institutional budget (not just the computing area’s budget) in measuring the “reasonableness” of required accommodations for accessibility.

This Act was the first act, which implemented the thought of providing accessibility aid to the people with disabilities.

Next the U.S. Rehabilitation Act of 1973 was extended with the Americans with Disabilities Act (ADA) in 1990. The ADA of 1990 [7] states:

- Extends the requirements of the Rehabilitation Act of 1973 to all public and commercial facilities, with few exceptions, not just those that receive federal funding.
- Requires that every institution receiving federal funds establish and maintain a plan of compliance.
- The Department of Justice stated that the ADA will cover government entities on the internet as well as those providers whose services are deemed to be public accommodations.

Finally, in 1998 an amendment was passed to Section 508 of the U.S. Rehabilitation Act of 1973 [7], [8]:
• States that federal agencies must ensure websites are accessible to employees and the public to the extent it does not pose an “undue burden” to the site owner.

• Provides accessibility standards, developed by the Architectural and Transportation Barriers Compliance Board (Access Board), for the web and many other areas of electronic and information technology [9].

• Standards for webpages were published Dec. 21, 2000.

• Enforcement date was set at June 21, 2001.

W3C launched the WAI in 1997 with endorsement by The White House and W3C members.

WAI develops guidelines for website development and other technical reports after evaluating the webpages. WAI pursues accessibility of the web through five primary activities [10]:

1. Ensuring that web technologies support accessibility
2. Developing guidelines for accessibility
3. Developing tools to evaluate and facilitate accessibility
4. Conducting education and outreach
5. Coordinating with research and development

To summarize, web accessibility enables diverse sets of users, regardless of disability or environmental constraints, can access material on a website. Different barriers hinder different groups and so a wide variety of barriers must be taken into account to ensure Content is accessible to the widest possible audience. The profile of web accessibility has been strengthened greatly in recent years. Society is starting to recognize that with an increasing number of online services becoming available, it is important that certain sections of society are not excluded.

On June 29, 2011, in the United States District Court Central District of California, Case No. CV 10-05810-DMG-FMO, Judge Dolly Gee certified a nationwide class of blind persons in a class action pending in Los Angeles, represented by counsel at Forizs and Dogali, P.A. The blind plaintiffs in the case do not seek money damages, but only seek an injunction requiring Disney to comply with the ADA by making its theme parks and websites accessible to persons with visual impairments [11].

The plaintiffs alleged and filed a legal case against that Disney’s websites do not meet ADA standards since they violated, the following [11]:
• Disney does not provide schedules, menus and maps in formats that are accessible to blind persons, such as in Braille, large print, or electronic form

• Disney does not accommodate the needs of blind persons during live parades and shows.

• Disney does not permit any discounted admission for sighted companions who must accompany and support blind persons in the theme parks.

• Disney’s websites do not follow ADA compliance norms, so that people with disabilities can also access information. Some of the accusations were:
  • Use of flash, which is not accessible to blind people,
  • Audio and video which cannot be turned off without a mouse, and
  • Audio and video drowns out the sound of the screen-reader.

Finally, by following ADA compliance and the rules laid out by WAI:

• It’s a win/win situation – all users benefit from a well-organized web site.

• For PDAs and for hand held devices like smart phone – accessible pages more easily convert to alternate platforms.

2.2 TECHNOLOGICAL SURVEY

Selection of the correct technology stack after the requirements are gathered is a critical and tedious process. This involves analyzing the requirements thoroughly and selecting the appropriate technology set, which matches our comfort level also. In this case, since I am well versed in PHP and once this tool goes live and it will be maintained and modified, if needed by the Art Lab IT consultant, who has a strong PHP background, PHP was selected as the programming language. In this chapter, I shall explain all the Open Source technologies, which has been used for the implementation of this application. The different Open Source technologies which have been used are:

• MySQL
• PHP
• HTML
• CSS
• JavaScript
• mPDF

MySQL provides an implementation of a SQL database very well suited for small to medium webpages. The database is free and Open Source with a commercial license
available (MySQL is now owned by Oracle after they bought Sun). Common applications for MySQL include PHP and java based web applications that require a DB storage backend, e.g. Dokuwiki, Joomla, xwiki, etc. Very many applications that use MySQL are geared towards the LAMP stack (Linux, Apache, MySQL, PHP). MySQL is usually used with two different storage engines, one is called MyISAM which does not support transactions and stores each table in a set of three files. The second is called InnoDB which supports transactions; this storage engine stores all data in a single set of bytes or uses one set of bytes per database directory [12].

MySQL has one major advantage: it is free. It is usually available on shared hosting packages and can be easily set up in a Linux, Unix or Windows environment. If a web application requires more than one database or requires load balancing, it is easy to set up many instances of the database requiring only the hardware costs, as opposed to commercial databases that would require a single license for each instance.

MySQL has some issues with stability and clustering; it is very difficult to install a consistent database cluster with MySQL with the regular version. Depending on the database storage, MySQL will support transactions or not, so the requirements of the application have to be taken into account when creating the database tables. For large, heavy loaded databases, it is a major operations problem that changing the database structure is only possible when locking the complete tables. This will mean that the database cannot be accessed during that operation so that this can only be done during low traffic times.

After the acquisition of Sun Microsystems by Oracle, some people are concerned about the future of the free version of MySQL, the original author of the database has started a project that uses a fork of MySQL to build a new database MariaDB (http://mariadb.org/), this is intended to be a compatible replacement of MySQL and plans to implement new storage engines later [12]. PHP is an Open Source server side programming language available free of cost that can be gotten easily from the market. Its coding style is quite easy to understand and it is very efficient on multi-platforms like Windows, Linux, and UNIX, and varieties of the latter two. It is a very flexible but powerful language, most suitable for developing dynamic webpages. Nowadays developing dynamic websites are in the huge demand due to its specific characteristics, like it automatically refreshes and does not need to make many changes manually [13]. In terms of advantage in running, PHP does not put
strain on servers. It uses its own inbuilt memory space that decreases the workload from the servers and the processing speed automatically enhances. Its script is optimized to make the server’s job easier, thus nowadays the uses of PHP is popular among programmers [14]. PHP also has an upper hand in running multimedia files, as PHP is not very dependent upon external plug-ins to run the programs. It is executed exclusively by the server and therefore needs nothing from the end users. PHP must be installed on the server and server parses the code at its source, executes and sends it properly send properly formatted HTML to the client computer. The user’s browser only sees code it understands. Whenever we begin any website, we consider several aspects of the site including website designing, web developing, web hosting, its maintenance, its flexibility, usability, scalability and how will it be easy to handle post development. After considering these aspects, if we decide to develop a site that looks attractive, smooth running, fast downloading, cost effective and easy to handle, we will definitely focus on PHP, one of the most popular web development scripts amongst the other available technologies like Java, ASP.Net, .JSP. Also one of the reasons PHP is selected and used by a lot of start up software companies is due to a budget factor, which is the most crucial part of the software development—especially for the small business users who wants to develop money making websites to earn thick profit in minimum investments. The limitation of budget is the biggest obstacle for small businesses. For those, PHP is the key solution as it is available at free of cost under PHP general public license and most of its associative required software, like MySQL, Text Editors and Apache Server, are also freely available, so it proves very cost effective for the developers. However, in terms of performance it is similarly efficient like other programming languages of same category.

Finally, summarizing the advantages of PHP [14]:

- PHP is accessible and available for free.
- There is a wealth of documentation available online, in many languages.
- PHP is a loosely typed language, which makes basic scripts work much faster and allows developers to concentrate less on design.
- PHP is flexible, using OOP or not.
- It is system independent, can run on Unix, Windows.
- Programmer of Java, Perl, C can find any parallels to ease transition to PHP. For example, JAVA developers who use MVC pattern for application development can also use MVC pattern in PHP by following Zend Framework in PHP.
But there are also a few disadvantages of PHP [15]:

- PHP executes more slowly than assembly, C and other complied languages.
- PHP is loosely typed language. And its for developers of all skill levels, this allows room for unexpected behavior due to programmer error which some languages might not permit.

HTML is the authoring language used to create documents on the web. HTML is used to define the structure and layout of a webpage, how a page looks and any special functions. HTML does this by using what are called tags that have attributes [16].

Origin of HTML dates back to 1997 and the credit goes to Tim Berners-Lee. It had 18 elements. This below piece of code is one of the oldest available HTML code developed. Incredibly this piece of HTML code even works today, this gives us an idea regarding the consistency in the structure and ability of the web browsers to interpret the code

```html
<title>Hypertext Links</title>
<h1>Links and Anchors</h1>
A link is the connection between one piece of
<a href=WhatIs.html>hypertext</a> and another
```

As the time evolved, better and sophisticated versions of HTML were developed. Cascading Style Sheets (CSS) is a style sheet language used for describing the presentation semantics of a document written in a markup language. Its most common application is to style webpages written in HTML and XHTM [17]. See Table 2.1 for evolution of HTML [18] and Table 2.2 for history of CSS [17] and Figure 2.2 for evolution of both HTML and CSS.

JavaScript is a prototype-based scripting language that is dynamic, weakly typed language. JavaScript was originally developed in Netscape, by Brendan Eich. JavaScript is used to make webpages more interactive and dynamic. JavaScript is a lightweight programming language and it is embedded directly into the HTML code. Many languages, especially Java, influenced JavaScript.

Advantages of JavaScript are [19]:

- HTML designers use JavaScript as one of the best programming tool that is very simple, syntax wise, and need not to have any previous strong programming skills because most of the HTML authors are not programmers. However, the syntax of JavaScript is so simple that most of the HTML authors easily play with it.
Table 2.1. Evolution of HTML

<table>
<thead>
<tr>
<th>Year</th>
<th>Version</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>HTML 1</td>
<td>HTML 1.0 was the first release of HTML to the world. Not many people were involved in website creation at the time, and the language was very limiting. Feature was limited to able to display some simple text onto the web.</td>
</tr>
<tr>
<td>November, 1995</td>
<td>HTML 2</td>
<td>HTML 2.0 included everything from the original 1.0 specification but added a few new features to the mix. HTML 2.0 was the standard for website design until January 1997 and defined many core HTML features for the first time.</td>
</tr>
<tr>
<td>January, 1997</td>
<td>HTML 3</td>
<td>More people were getting into the HTML development around now, and they wanted to enhance the look of their sites. This also led to Browser wars, since some of the HTML tags were not compatible with other browsers. A HTML working group, led by a man named Dave Raggett introduced a new HTML draft, HTML 3.0. It included many new and improved abilities for HTML, and promised far more powerful opportunities for webmasters to design their page. But browsers were not yet to the speed, this alerted the developers to modularize their code.</td>
</tr>
<tr>
<td>December, 1997</td>
<td>HTML 3.2</td>
<td>The browser specific tags kept coming, and it was high time for a standard to be established W3C was founded in 1994 to standardize the language.</td>
</tr>
<tr>
<td>December, 1999</td>
<td>HTML 4.0.1</td>
<td>HTML 4.0 was recommended by the W3C in December ‘97 and became the official standard in April 1998. Microsoft in their Internet Explorer browser had excellent support for almost all of the new tags and attributes. Once HTML 4.0 had been out for a little while, the documentation was revised and corrected in a few minor ways and was entitled HTML 4.01; the final version of the specification.</td>
</tr>
<tr>
<td>January, 2000</td>
<td>XHTML 1.0.0</td>
<td>Close to the beginning of the 21st century the W3C issued their specifications of XHTML 1.0 as a recommendation. XHTML it is an entirely new branch of HTML, incorporating the rigors of XML, so that code must be properly written if it is to work once it reaches the reader’s browser. There weren’t many new or deprecated tags and attributes in XHTML, but some things changed with a view of increased accessibility and functionality. It was mainly just a new set of coding rules.</td>
</tr>
<tr>
<td>January, 2008</td>
<td>HTML 5</td>
<td>It was around this time that a bunch of pragmatic web technology fans, browser programmers and specification writers started building something of their own, outside of the usual W3C procedures. They called themselves the Web Hypertext Application Technology Working Group (WHATWG), and developed a new spec. After some soul-searching, the W3C decided that HTML was still the future of the web. And HTML5 became the new specification that everyone’s effort should be poured into. HTML5 is designed for the web, both now and in the future. This is the specification that we will be working with for the next decade at least, so the process of its development is relatively slow and considered. Many parts will be familiar, but there are also plenty of new elements, attributes and abilities to get excited about. You can check the latest version of the spec if you want all the detail. A full tutorial on HTML Source about the changes in HTML5 is forthcoming.</td>
</tr>
</tbody>
</table>

Table 2.2. History of CSS

<table>
<thead>
<tr>
<th>Year</th>
<th>Version</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>December, 1996</td>
<td>1 ~</td>
<td>Font properties such as typeface and emphasis</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>Color of text, backgrounds, and other elements</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>Text attributes such as spacing between words, letters, and lines of text</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>Alignment of text, images, tables and other elements</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>Margin, border, padding, and positioning for most elements</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>Unique identification and generic classification of groups of attributes</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>The W3C no longer maintains the CSS Recommendation.</td>
</tr>
<tr>
<td>May, 1998</td>
<td>2 ~</td>
<td>A superset of CSS 1, CSS 2 includes a number of new capabilities like</td>
</tr>
<tr>
<td></td>
<td></td>
<td>absolute, relative, and fixed positioning of elements and z-index.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Concept of media types, support for aural style sheets and bidirectional text,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and new font properties such as shadows.</td>
</tr>
<tr>
<td></td>
<td>~</td>
<td>The W3C no longer maintains the CSS 2 recommendation.</td>
</tr>
<tr>
<td>December, 2010</td>
<td>2.1 ~</td>
<td>Fixed errors in CSS 2, ~ Removes poorly supported or not fully</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interoperable features and adds already-implemented browser</td>
</tr>
<tr>
<td></td>
<td></td>
<td>extensions to the specification.</td>
</tr>
<tr>
<td>June, 1999</td>
<td>3 ~</td>
<td>Unlike CSS 2, which is a large single specification defining various</td>
</tr>
<tr>
<td></td>
<td></td>
<td>features, CSS 3 is divided into several separate documents called</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“modules”. ~ Each module adds new capabilities or extends features</td>
</tr>
<tr>
<td></td>
<td></td>
<td>defined in CSS 2, over preserving backward compatibility.</td>
</tr>
<tr>
<td>Sep, 2009</td>
<td>4 ~</td>
<td>Under review. Currently not supported in any browser.</td>
</tr>
</tbody>
</table>


- JavaScript can be used to make HTML pages more dynamic. All plain HTML files are static.
- JavaScript has the ability to react to various events, which is not possible through HTML. When a user clicks a button, an action event can be generated with the help of JavaScript.
- JavaScript can be embedded easily into an HTML document and JavaScript easily reads and can write HTML elements as well.
- To detect the browser of the visitor, we can use JavaScript and can load the page according to the requirement of that browser.
- Cookies can be created using JavaScript.
- You can add various other features like user validation, authentication, etc. by the use of JavaScript.
- JavaScript is embedded directly in to the markup code or the HTML code.
- JavaScript is an interpreted language in which scripts are directly executed without preliminary compilation.
- We are not required to purchase a license in order to use JavaScript; it is Open Source.
This application works even if the JavaScript option is turned off in the browser. Then the application would adjust its behavior to the browser’s nature.

There are quite a few Open Source PDF conversion libraries available in the web. These libraries basically generate PDF files from HTML.

mPDF is one such PHP class, which generates PDF files from UTF-8 encoded HTML. It is based on FPDF and HTML2FPDF with a number of enhancements. Main features of this library are [20]:

- Accepts UTF-8 encoded HTML
- Supports almost all languages including RTL (Arabic and Hebrew), and CJK - (chinese-japanese-korean)
- Bookmarks
- CSS style sheets
• Word spacing and character spacing for justification
• Nested block-level elements (e.g. P, DIV) including margins, borders, padding, line-height, background colors, etc.
• Support (partial) for floating and fixed-position block-elements
• Page layout and orientation
• Text-justification and hyphenation
• Page numbering
• Odd and even paging with mirrored margins
• Page headers and footers
• Columns
• Tables - nested tables, rotated, or auto sized to fit on a page
• Table of contents
• Index
• Watermarks
• Images in JPG, GIF, PNG, SVG, BMP or WMF format
• Password protection
• Annotations
• Barcodes (EAN13, UPC-A/E, Code 11, 39, 93, 128, Codabar, MSI, IMB, Planet, Postnet, RM4SCC, etc.)
• Import another PDF file and use as a template
• Joined Arabic text - support for Arabic, Persian, Urdu, Sindhi and Pashto
• Support for India languages (incl. consonant conjuncts) - Gujarati, Devanagari, Bengali, Tamil, Malayalam, Oriya, Kannada, Telugu
• Embedded font subsets

2.3 ORIGIN AND OVERVIEW OF WEB BASED CONTENT MANAGEMENT SYSTEMS

The pre-millennial web was characterized by highly manual approaches to maintenance; the successful and substantial postmillennial web has had a significant automation. One process, which has heavily contributed to this transformation, is the CMS [21]. Features, which are the core concepts that defines a CMS:

• Repository in file system or database
• Separation of content from presentation
• WYSIWYG editing
• Workflow - automation of tasking sequences and business rules
• Link management
• Metadata - for a search engine and advanced navigation
• Multichannel delivery (print, PDF, PDAs, cell phones, etc.)
• Personalization
• Separation of creation and content management from delivery

All these features shall now be now explained, with regard to my application.

2.3.1 Repository in File System or Database

Uploaded images are physically stored on the server; the image path (the image index) is stored in the database. Check-In-Check-Out: checking out a piece of content puts a file lock on the content so someone else does not change it before the current user’s changes are saved. A good system tells a new user who has the content, and provides details to contact the current user to release it. They also have administrative override privileges to seize content. Metadata search: metadata is descriptive data that is attached to an object (element) that allows a piece of content to be more accurately indexed, retrieved, or processed. It describes the content, and might provide optional information like artist name, genre, and date of creation for the search engine to retrieve the associated data. This metadata is usually stored in a relational database or an object oriented database (the CMS Repository). It can be retrieved thorough different queries.

2.3.2 Multichannel Delivery

Multichannel Publishing means different versions of the content can be formatted for delivery in different physical “channels” like the web (HTML and PDF), traditional print, wireless handheld devices, and cell phones. Adobe calls it “network publishing.” In this current project the scope of publishing is limited to web and PDF. Separation of content from presentation: The user of the application does not have to worry about how the images are getting resized or how the Image gets converted into PDF format. See Table 2.3 for an overview of CMS [21].
Table 2.3. CMS Explained

<table>
<thead>
<tr>
<th>Process /Benefit</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowing reuse of information through integration of data from different sources.</td>
<td>The CMS welcome page for Art lab can draw together, information from different sources like student, faculty privileges, image database, slide shows information.</td>
</tr>
<tr>
<td>Permitting the efficient re-purposing of information</td>
<td>The same image can be rendered as a PDF or as image embedded in HTM</td>
</tr>
<tr>
<td>Ensuring presentational consistency by separating the design of webpages from content they display</td>
<td>Faculty or Admin will be provided with a template to upload images, update metadata, info.</td>
</tr>
<tr>
<td>De-skilling the task of putting information on the web.</td>
<td>Which reduces the task of simple browse and upload images for faculty and download for students</td>
</tr>
<tr>
<td>Facilitating good information management practice so that appropriate metadata are captured at the time of content creation and modification</td>
<td>Every time a image is uploaded, the created by metadata field should be populated, the date and time created should be provided, this aids in metadata searching and also acts storehouse of information, to track events</td>
</tr>
</tbody>
</table>


2.4 ORIGIN AND ADVANTAGES OF PDF

Portable Document Format (PDF) is a kind of file format introduced by Adobe and is one of the most popular file format, used by computer users worldwide. Most computer users are likely to have used PDF documents in the form of e-books, as a part of documentation file for software purchased, while downloading some registration form from a website. This would trigger a question and analysis regarding what are these PDF documents and why are they being used widely despite the presence of other software available like Open Office or Microsoft Office. Advantages of using PDF files [22]:

- Fast and simple to create: PDF files are very easy to create and virtually take no time at all. It is as easy as clicking a button with your mouse. Most importantly PDF is designed to keep up with today’s fast-paced and highly demanding work world. PSD files can be read through free software programs like PDF viewer. Creating a PDF is meant to take up almost no time at all.

- Security: Security has played a major role in the popularity of PDF documents. If you are worried about sending your documents over internet or circulate it in the office, no need to worry with PDF documents. All PDF files support data encryption, hence providing an easy and secure way of data transfer and sharing.

- Capability to compress large files: Imagine storing and transferring large volume of heavy files. PDF has made life lot easier.
• Portability: All PDF files exist independently from the operating system, hardware, software that is being used to view, edit or convert PDF files. It not only makes it easy to transfer files among computers with different hardware and software architecture. Using PDF files also make sure that the document you send reach the destination clearly without any error.

• Printability: When we convert other formats such as images to PDF, it can help by saving the space. In addition, the PDF format can be easily printed with its fast open speed and quickly printing processing, which also becomes a factor to vote for its superb portability.

• Security and Preservation: PDF format is a read-only format and readers are not allowed to modify, edit, copy, print and even read at random with security restrictions. Besides, due to the stability of the formatting and contents, the document can maximally preserve the authors’ original intention and protect their content.

• Independence and Multi-Platforms: We can open PDF files under any applications, software, hardware, and operation systems at your will. The format is independent of the platforms and can always preserve the fixed-layout and styles including graphics, texts, fonts, spaces, tables, and other information reliably.

These were the advantages of using PDF files. But difficulty or inability to edit PDF files is a cause of major concern among different users. By its nature, PDF is not an editable document format. And also these are some of the interesting feature, which distinguishes PDF from DOC format and makes its more preferable over DOC format.

Scenarios when it is suitable to create PDF:

• Intend to print your documents
• Want to read eBooks on large screen
• Prepare to back up your document files
• Eager to share/deliver your writing with/to other people
• Protect your copyright by add PDF password or DRM
CHAPTER 3

IMPLEMENTATION

3.1 ARCHITECTURE

Figure 3.1 shows high-level architecture of the system. The CMS is built on top of PHP, which sits on top of Apache.
My role in this thesis can be attributed to these sections:

- Selection and implementation of design pattern
- Image processing
- Metadata search
- HTML to PDF conversion

### 3.2 Model View Controller Pattern

The site architecture was built using the model view controller (MVC) pattern. The MVC pattern is the most used pattern for today’s world web applications. It was used for the first time in Smalltalk and then adopted and popularized by Java. The MVC pattern separates an application in three modules, model, view and controller:

- The model is responsible to manage the data; it stores and retrieves entities used by an application, usually from a database, and contains the logic implemented by the application [23].
- The view (presentation) is responsible to display the data provided by the model in a specific format. It has a similar usage with the template modules present in some popular web applications.
- The controller allows the model and view layers to work together. The controller receives a request from the client; invokes the model to perform the requested operations and send the data to the view. The view formats the data to be presented to the user, in a web application as an HTML output.

See Figure 3.2 for general architecture of MVC [23].

![MVC collaboration diagram](http://PHP-html.net/tutorials/model-view-controller-in-PHP/)

3.2.1 Controller

The controller is the first thing which takes a request, parses it, initializes and invokes the model and takes the model response and send it to the presentation layer. It is practically the connection /coordinator between the model and the view, a small framework where model and view are plugged in. In our native PHP implementation the controller is implemented by only one class, named “controller.” The application entry point will be index.php. The index PHP file will delegate all the requests to the controller. MVC code snippet of controller:

```php
// index.PHP file
include_once("controller/Controller.PHP");
$controller = new Controller();
$controller->invoke();
```

3.2.2 Model and Entity Classes

The model represents the data and the logic of an application, what many calls business logic. Usually, it is responsible for:

- Storing, deleting, updating the application data. Generally it includes the database operations, but implementing the same operations invoking external web services or APIs is not an unusual at all.
- Encapsulating the application logic. This is the layer that should implement all the logic of the application. The most common mistakes are to implement application logic operations inside the controller or the view (presentation) layer.

MVC code snippet of model view (presentation) is as follows:

```php
include_once("model/Book.PHP");
class Model {
    public function getBookList()
    {
        // Here goes some hardcoded values to simulate the database
        return array(
            "Moonwalker" => new Book("Moonwalker", "J. Walker", "")
            "PHP for Dummies" => new Book("PHP for Dummies", "Some Smart Guy", "")
        );
    }
    public function getBook($title)
    {
```
// We use the previous function to get all the books and then we return the requested one.
// in a real life scenario this will be done through a db select command
getAllBooks = $this->getBookList();
return $allBooks[$title];
}
}

The view (presentation layer) is responsible for formatting the data received from the model in a form accessible to the user. The data can come in different formats from the model: simple objects (sometimes called value objects), XML structures, JSON.

The view should not be confused to the template mechanism sometimes they work in the same manner and address similar issues. Both will reduce the dependency of the presentation layer of from rest of the system and separates the presentation elements (HTML) from the code.

The controller delegates the data from the model to a specific view element, usually associated to the main entity in the model. For example the operation “display account” will be associated to a “display account” view. The view layer can use a template system to render the HTML pages. The template mechanism can reuse specific parts of the page: header, menus, footer, lists and tables. MVC code snippet of view is as follows:

```html
<html>
<head></head>
<body>
<?PHP
    echo ‘Title:’ . $book->title . ‘<br/>‘;
    echo ‘Author:’ . $book->author . ‘<br/>‘;
    echo ‘Description:’ . $book->description . ‘<br/>‘;
?>
</body>
</html>
```

In this application, everything inside the ‘lib’ directory implements the ‘model’. This includes the classes that interface with the database, and classes that fetch and process the image objects. The classes in ‘lib/us’ implement the views. These classes are responsible for formatting the data and presenting it in HTML format to the controller. The classes in ‘lib/controllers’ implement the controller. These classes are responsible for handling the GET and POST HTTP request and handling the site-wide common functionalities such as authentication and logging.
3.3 IMAGE PROCESSING

The application supports two formats of images in our application - JPEG and PNG format. The maximum size of the image is limited to 5 MB. The image uploader strictly enforces these constraints during the upload process.

Every image uploaded to the application is processed and re-sized to optimize the bytes transferred over the network. Most often whenever we display the images in the application it is sufficient to render the smaller version of the image instead of the original high-resolution image. Since we cannot do the re-sizing every time on the fly, we do a one-time re-sizing of the images when they are uploaded. Resizing logic has been embedded into the code. The images are re-sized into three different resolutions:

- **Thumbnails:** This is a 100 px width and 100 px height image. It is used whenever we want to render the image objects in a group. For example, managing the entire image or to show in a presentation.

- **Medium:** This is 800 px width and 600 px height image. This size is most ideal in the cases when a larger view of the image is required but not necessarily a higher quality image. This is used in the context of viewing and editing an image object.

- **Large:** This is 1024 px width and 960 px height image. This resolution is desired when a higher quality image is required to render. We use the images with this resolution in the light box view in the presentation objects.

Apart from these resolutions, we also keep the original high-resolution image, which we use in the PDF generation since an image with higher PPI (Pixel per Inch) is required to have a good print of the flashcard presentations. See Figure 3.3 for screenshot of image management page.

3.4 SEARCH FUNCTIONALITY

The application provides the functionality of searching the image objects by their metadata. The metadata is indexed and stored in the DB. The keywords given to the search system are used to construct an SQL query to perform the search. The constructed SQL query does a table join on the metadata and the resource table to quickly search through the metadata of relevant image objects. There are two types of search functionality provided:

- **Simple Search:** In this mode the faculty can enter one or more keywords into the search field and the system returns the images containing any of those keywords. In this mode the keywords are simply ‘OR’ while constructing the SQL query.
Advanced Search: In this mode the faculty can search by keywords specific to the metadata field. In this mode the keywords are ‘OR’ within the metadata field and the result set is intersected to produce the final output.

A shortcut to create a presentation out of the images from a search result is provided to the faculties. This feature would be useful to quickly create a presentation of related image objects. See Figure 3.4 for a screenshot of search functionality.

### 3.5 HTML TO PDF CONVERSION

As explained earlier mPDF library has been used inside the PHP code, which allows the conversion of images embedded in HTML to be PDF. This Open Source library has been integrated with MVC code, so that when the controller get the request for the PDF conversion it routes the request to this class, which internally converts the HTML images to PDF format.
3.6 ENVIRONMENTAL SETUP FOR THE APPLICATION

The hardware setting of the application, which was used for development follows. Code has been developed incrementally on this server, so environmental issues will not be faced once the application goes live.

3.6.1 Production Hardware Overview

The specification of the production hardware are:

- Model Name: Xserve
- Processor Name: Quad-Core Intel Xeon
- Processor Speed: 2.8 GHz
- Number Of Processors: 1
- Total Number Of Cores: 4
- L2 Cache: 12 MB
- Memory: 4 GB
- Bus Speed: 1.6 GHz
- OS: Mac OS X Server 10.6.8 (Darwin core)
- Apache 2.2.1, PHP Version 5.3.8, MySQL version 5.0.8
3.6.2 Hardware Overview

Development server it is currently on:

- Model Name: Xserve
- Processor Name: Dual-Core Intel Xeon
- Processor Speed: 2 GHz
- Number Of Processors: 2
- Total Number Of Cores: 4
- L2 Cache (per processor): 4 MB
- Memory: 4 GB
- Bus Speed: 1.33 GHz
- OS: Mac OS X Server 10.6.8 (Darwin core)
- Apache 2.2.1, PHP Version 5.3.8, MySQL version 5.0.8
CHAPTER 4

TESTING

Software testing is an essential step in order to verify if the software is working as expected and to ensure that it is behaving the same for even corner cases. The two major approaches of software testing are manual software testing and automated software testing. Manual software testing means it is being done by a person, i.e. a tester runs the software for errors. In case of automated testing the automated scripts, regression tests, predefined use cases test the flow of the application.

A few testing types and how these testing types [24], were adapted in my application are as follows:

- **Black box testing**: Internal system design is not considered in this type of testing. Tests are based on requirements and functionality.
  
  **Scenario**: End user who has no knowledge of how the application was built was given an functional overview about the application and was asked to upload, edit metadata and create slides.

- **White box testing**: This testing is based on knowledge of the internal logic of an application’s code. Also known as Glass Box Testing. Internal software and code working should be known for this type of testing. Tests are based on coverage of code statements, branches, paths, and conditions.
  
  **Scenario**: Checking if the images were being resized as expected.

- **Unit testing**: Testing of individual software components or modules. Typically done by the programmer and not by testers, as it requires detailed knowledge of the internal program design and code.
  
  **Scenario**: Test the admin features of the application by creating users, i.e. student roles, faculty roles and deleting users.

- **Incremental integration testing**: Bottom up approach for testing, i.e. continuous testing of an application as new functionality is added. Application functionality and modules should be independent enough to test separately.
  
  **Scenario**: Once the admin panel was built, PDF conversion process was developed and tested simultaneously, to ensure the images were getting converted into PDF format.

- **Integration testing**: Testing of integrated modules to verify combined functionality after integration. Modules are typically code modules, individual applications, client
and server applications on a network, etc. This type of testing is especially relevant to client/server and distributed systems.

- **Functional testing:** This type of testing ignores the internal parts and focus on whether the output is as per requirement or not. It is similar to Black-box type testing geared to functional requirements of an application.

- **End-to-end testing:** Involves testing of a complete application environment in a situation that mimics real-world use, such as interacting with a database, using network communications, or interacting with other hardware, applications, or systems if appropriate.

  **Scenario:** Testing the entire application in detail, checking if the navigation between pages is working as expected, search tabs are returning results as expected, image upload process is working as expected.

- **Sanity testing:** Testing to determine if a new software version is performing well enough to accept it for a major testing effort. If application is crashing for initial use then system is not stable enough for further testing and build or application is assigned to fix.

  **Scenario:** Initially the metadata search was working for only one type of input criteria, and later the code was updated, so that multiple values can be passed in the input fields in order to retrieve the associated info.

- **Stress testing:** System is stressed beyond its specifications to check how and when it fails. Performed under heavy load like putting large number beyond storage capacity, complex database queries, and continuous input to system or database load.

  **Scenario:** Trying to upload images, which are bigger in size [greater than 5 MB] and verifying if the image path is being persisted in the database. Trying to search the images by passing in a variety of combinations of metadata input.

- **Security testing:** Can system be penetrated by hacking way. Testing how well the system protects against unauthorized internal or external access. Checked if system, database is safe from external attacks.

  **Scenario:** Users cannot capture and reply the old requests to cause malicious attacks, such scenarios are handled by unique CSRF tokens, embedded in the request, response.

- **Alpha testing:** In house virtual user environment can be created for this type of testing. Testing is done at the end of development. Still minor design changes may be made as a result of such testing.

  **Scenario:** Some of the UI related changes were identified during this phase, such as re ordering of tabs, and alignment of input field for text box.
CHAPTER 5

OBSTACLES AND RESULTS

During the development of this application few roadblocks were encountered. I have listed a few of them below.

5.1 HTML TO PDF CONVERSION

There are only a few Open Source and free libraries available, in order to convert the HTML to PDF. Also after analyzing and selecting mPDF as conversion tool, integration and usage of this class into the MVC code was a challenge. Since, in MVC pattern the control takes the requests and internally the request had to be routed to class implementing this API.

5.2 IMAGE RESIZE

As mentioned earlier, this application resizes the images into thumbnail, medium and large sizes. But this had to be accomplished in such a way that while resizing the images to thumbnail, the clarity of the images are not affected drastically and each image being considered a resource, upon upload each image had to be indexed and all the three versions image had to be persisted on the server and only the image location was persisted in the database.

5.3 METADATA SEARCH

This metadata search tab can search the images even if the complete metadata information is not provided. Like SQL clause has been used in the SQL queries and also due to functionality of the information will be retrieved based on multiple criteria passed in the request, complex queries were constructed to achieve it.

The results of the application can be summarized as follows:

- **Metadata Search Output**: The queries are fine tuned such that images are returned back between 0.35 – 0.75 milliseconds, based on the search criteria.
- **Image and Metadata Upload**: The image upload process takes about a second and it resizes the images into 3 different sizes and stores the actual images on the server and persists the handle / index of the image, i.e. the location of the image in the database.
Along with the image, a user can also insert the metadata information that better describes the image. This metadata information associated with an image is stored in the database.

- **PDF Slideshow:** A faculty can select images by searching through the metadata search tab and create a custom slide show and then publish it to the students. Currently there is no limit to the number of images being added to a slideshow.

- **Download PDF Slideshow:** Students can easily download the PDF slideshow uploaded by the faculty, so that it can be accessed online and offline, even without connecting to the internet. Since most of the smart phones have inbuilt PDF readers, this site is also operable on a smart phone.
CHAPTER 6

FUTURE ENHANCEMENTS

There is always a scope and room for improvement for any software application. With the current set of requirements and also due to time constraint, currently the application is working as expected and will cater to the needs of the Art Lab, SDSU.

One of the major improvisations, which can be done to this CMS, is to achieve language localization. Currently, the system is fully functional with English being the medium of language display. In the future the application can be extended such that it supports and displays languages like Hindi, Spanish, German, and French, etc.

Currently, images are being uploaded, but in the future, the application can be extended to upload any video or audio, any such multimedia related data.

Also another area where this application can be improvised is by providing the check-in check-out features, version control and versioning, with scheduling. So that if an image is being worked by a person, it has to be locked and only when its released another person with the appropriate credential gets to work on the image.
CHAPTER 7

CONCLUSION

In this research I have explored CMS in detail. Analysis was done on how the CMS has evolved, since its initial origin. Also, the importance and origin of ADA compliance was covered. Integration of user accessibility features with the user related functionality at an application level was achieved.

CMS for Art Lab was designed and implemented to improve the current process of image and metadata upload process. With this system now in place, image and metadata management will be a streamlined and robust process.

ADA compliance regulations have been studied and its implications have been thoroughly understood. And these ADA compliance rules were incorporated into this application, so that people with disabilities can also use this web application without any shortcomings.

Analysis on PDF and its benefits has been successfully conducted. PDF generation process has been reviewed, and this feature has been integrated with the application.
BIBLIOGRAPHY


