Creating Virtual Exhibitions from an XML-based Digital Archive

Johanna Chua Lim and Schubert Foo(*)
Division of Information Studies, School of Communication & Information, Nanyang Technology University, 31 Nanyang Link, Singapore 637718
Email: assfoo@ntu.edu.sg(*)

Abstract

A virtual exhibition system was designed and implemented to interface with the National Archives of Singapore digital archive to support the creation of online virtual exhibitions. The XML-based digital archive provides different artifact types that form the contents in the exhibition through the reference and reuse model. Artifacts and exhibitions are endowed with rich metadata that include the Dublin Core elements to support enhanced search support for field and free text searches. An authoring tool using a grid-layout approach is used to define and layout the exhibition contents. XML’s Cascading Style Sheet and Extended Style Sheets are then selected and applied to the XML documents from the digital archive to yield the final exhibition in HTML format. By supporting the use of information layering in the descriptors of the artifacts and the application of different style sheets, it becomes possible to create multiple versions of the same exhibition that vary in content, layout and interface so that it can be used to meet the diverse information needs of a range of different user communities.

Keywords:
XML-based digital archive, virtual exhibitions, reusable digital artifacts, metadata, National Archives of Singapore

Introduction

The National Archives of Singapore (NAS) is “responsible for the collection and management of the nation’s (Singapore) public and private historical records, in various formats” (“National Archives of Singapore”, n.d.). Over the years, NAS have created a digital archive comprising photographs, audio visual, cartographic and architectural plans, government and private records, posters, speeches, and so on. Many different applications and services can be created and developed to interface with this architecture to serve the information needs of its different user communities. Each different digital artifact is endowed with a set of rich metadata that includes support for the Dublin Core metadata set to support multiple access and differing uses by these applications.
Not only should an artifact be captured and stored with the best possible technical solution available but information pertaining to the artifact must also be complete and well organized for it to retain its value for future generations. As such, the artifact’s information collection process must be done accurately and thoroughly to avoid losing important information so that usefulness of the metadata can be maximised and exploited in future.

This paper reports on the design and development of virtual exhibitions using this digital archive as one example of the many potential applications that can be developed in future. The choice of the virtual exhibition as a typical application of the system is motivated by a number of reasons. Virtual exhibits are used by digital museums to provide varied museum visitors access to their collections. Exhibitions are generally thematic in nature and draw its resources from the rich collection of materials in the archive to create exhibitions to inform and educate the public. The availability of digital artifacts makes it possible to display content rich virtual (or online) exhibitions, which increases the value of museum information by providing its visitors with artifact’s acquisition information and detailed information about the artifact when asked (Callery & Thibadeau, 2000).

Virtual exhibits can be used to promote museum education through its interactive multimedia technology, permits multiple visitors simultaneous access to a single artifact, and provide quick and enriching browsing of the collection (Takahashi, Kushida, Hong, Rieger, Martin, Gay, Sugita, Kurita, Reeve & Loverance, 1998). Virtual exhibitions are not bounded by physical space and time unlike exhibits in traditional museums. Not all artifacts in the collection can be exhibited in a traditional museum, as this has to give way to other new exhibits. However, this is not the case with virtual exhibitions since collections can be made totally accessible; exhibitions maintained for future virtual visits, and presented to allow visitors have access to additional information about the exhibition or artifact. The use of high-resolution graphics, slide shows, audio narrative, video clippings in a virtual exhibit enhances the educational experience of visitors especially children. It will not only help ordinary visitors understand the exhibit but will also allow people of various kinds of physical disabilities to enjoy the benefit of museums (Koshizuka & Sakamura, 2000). However, it is acknowledged that multimedia presentation requires high bandwidth, which may not be available to certain parts of the world thereby depriving potential visitors’ access to the exhibit. Currently, website contents implemented in multimedia are generally not indexed and searchable by search engines. This limitation prohibits users from looking for specific exhibition pages that satisfies their information needs. They have to rely largely on browsing to interact with virtual exhibitions.

The study therefore aims to demonstrate how a XML-based digital archive can be used in the development of the Web-based Virtual Exhibition System (VES) for NAS by defining a rich layered-set of metadata of image and text artifacts, developing an efficient method of creating virtual exhibitions, implementing field-based and free-text based search and retrieval of the exhibition and artifact information, applying different design and layout
templates and styles to the same exhibition content to create multiple versions of exhibition in terms of content and style to cater to visitors of different age groups and information needs. With VES in place, it becomes a simpler task to create new Web-based virtual exhibitions in HTML format that can be furthered enhanced manually using traditional HTML authoring tools, if necessary. In this study, we will limit the exhibition content to include only text and photograph digital artifacts.

The rest of the paper is organized as follows. An introduction to the NAS digital archive followed by the descriptions of the VES main design issues and components that include the use of XSL (part of the XML family of tools) to create different look-and-feel exhibitions, metadata design and implementation, VES authoring and searching tools. The paper concludes with the salient points of the study and proposes a number of areas for future work.

**NAS Digital Archive Architecture**

The NAS digital archive 3-tiered layered XML-based architecture comprising the infrastructure layer, application layer and the end user layer is shown in Figure 1. The infrastructure layer consists of the artifact repositories with associated metadata information, indexing and retrieval engines, and repository gateway. The artifact repositories store the various types of NAS artifacts that can be accessed and manipulated by the different applications. Applications may also have its own associated metadata and specialized content to form application-specific repositories. An example of this is VES which contains its own set of exhibition metadata stored in the exhibition repository. The indexing and retrieval engine that takes the form of an XML server support the search and retrieval of artifacts stored in their respective repositories. The server also doubles up as a link to the different artifact repositories. The repository gateway serves as the communication link between the application layer and the infrastructure layer by translating the application layer’s request into the XML Server’s syntax and translating the XML Server’s response into a form the application layer can process. It basically comprises a number of APIs (Application Programming Interface) to support a set of primitive and higher-end operations that are needed by the various applications.

The application layer provides the functionality of the digital archive. It consists of the different application systems that administer the repositories, retrieves and process the data in the repositories to meet the various objectives of these applications. The web server also resides in the application layer and serves as the communication link between the users and the applications.

**Virtual Exhibition System Architecture and Design**

Virtual exhibitions are collections of Web pages revolving around a topic. It may be subdivided into different parts where each part corresponds to a sub-topic in the exhibition. Each exhibition part contains one or more pages. Each page of a virtual exhibition contains references to other pages, references to artifacts and local Web page information. The artifacts that comprise the content of the exhibition’s page may be
digital photographs, texts and other multimedia resources that are stored in the individual object repositories. Local web page information refers to the exhibition’s page content that is not stored in the repository but is used in the exhibitions such as Web banners and navigational buttons. This information resides in the Web server hosting the exhibition.

The VES architecture is based on the reference and reuse model of creating virtual exhibitions so that multiple copies of the same artifact are not copied and stored separately in different exhibitions (Lim, 2002). VES provides three main tools for the users to create and view virtual exhibitions as shown in Figure 2.
Figure 2. VES processes and architecture
**Artifact Definition**

Virtual exhibitions are conceived, designed, authored and implemented by virtual exhibition authors. An exhibition author starts off by doing the necessary research on the exhibition theme, forms the exhibition concept, and selects the various objects (i.e. content) to be used for the exhibition. If the objects already exist in the repository, they can be searched and retrieved, or browsed by the author for selection. If new objects are used for the first time, it must first be inserted into their respective repositories.

In VES, an *Artifact Metadata Tool* is provided for the author to create and store the different artifact objects used in the exhibition. Each of these created metadata will have a common set of Dublin Core metadata elements that serve as the foundation metadata elements in addition to other specialized metadata to fully describe different artifact types. Different repositories are used to store each of these artifact types. Metadata stored in the repositories will be indexed via the XML server to facilitate the searching and retrieval of artifacts and exhibition data by the information retrieval engine that is actually part of the functionality of the XML server. Copies of each artifact metadata are also stored in the web server’s local drive for back-up purposes, repopulating the repositories and to provide direct access to the metadata without the need to retrieve it from the repositories. Each artifact’s metadata makes reference to the actual digital artifact stored in the web server’s local drive. These actual artifacts are reference to when the HTML exhibition page is created, accessed and viewed eliminating the need to retrieve objects from the digital archive repositories each time an exhibition is accessed. However, with text artifact, its metadata includes its actual content to facilitate indexing, searching and retrieval.

**Virtual Exhibition Creation**

After determining the content of the exhibition, the author organizes and lays out the various exhibition pages and defines the objects in each page using VES’s *Virtual Exhibition Authoring Tool*. This tool is used to create the exhibition metadata and HTML exhibition pages. The created exhibition metadata is stored in the exhibition repository and the web server’s local drive just like the created artifact metadata. The exhibition page content includes references to the different artifacts. Artifact metadata are retrieved from the artifact repository based on their reference ID, and is incorporated in the page content together with the actual artifact stored in the web server’s local drive.

In VES, the exhibition content is kept separate from page formatting and layout. Exhibitions content is described using XML while formatting and layout templates in the form of Cascading Style Sheet (CSS) and Extended Style Sheet (XSL) are used. CSS provide display related information such as table height and width, font definitions, colours for web pages, and so on. XSL is used to format the artifact and exhibition’s XML documents into HTML pages for display in Web browsers. Formatting is also applied to each retrieved artifact, which provides exhibition visitors detailed information of an artifact upon demand.
The separation of content and format facilitates the reusability of exhibition’s content. Multiple exhibitions may make reference to the same content and since content is separated from the formatting information, the reused content remains intact. New formatting templates may be applied to the reused content creating exhibitions with the same or similar content but different presentation style. Information layering is also possible with the use of XSL templates. The XSL template applied to the XML document may determine which specific information to pull and display based on the needs of the visitors thereby creating exhibitions with different levels of information been displayed. For instance, an original photograph can have different scanned sizes, different resolutions to support zooming or thumbnail presentation, or digitally-manipulated, etc. These basically form a series of digital objects of an original photograph that share the same metadata and stored in the same XML file. Likewise, a text object may be described in different degrees of granularity or written using different styles for different age groups or interests of users. Such layering of information provides the exhibition author with a variety of choices to determine the content to be displayed. This benefit may be extended for information exchange where the necessary information in whole or in part is transmitted to the other party without the need to modify the content.

**Searching Virtual Exhibitions**

In interacting with the virtual exhibits, visitors can utilize VES’s *Exhibition Search Tool* to search for information in these exhibitions. Although this tool appears integrated with the system, it is actually a stand-alone application that allows searching for all the indexed information via the XML server. Searches can be done for specific occurrence of particular artifacts in the exhibitions, as well as exhibition and artifact related information that are stored in the repositories. The matching search results are retrieved from the different repositories and formatted for display. The search output makes direct reference to the exhibition’s HTML pages and artifact detailed HTML pages. The later caters to the same artifact being used across a number of different exhibitions.

**Metadata Design**

Metadata is data about data and in an archival setting, a “metadata accompanies and makes reference to each digital object and provides associated descriptive, structural, administrative rights management, and other kinds of information” (Day, 2001). So even when an individual, computer systems and information standards under which the metadata was created no longer exist, the information referring to each digital object can still be accessed and used. Thus, metadata is designed to support the retrieval and display of resources (Vercoustre & Paradis, 1999).

We adopted the Dublin Core (DC) Version 1.1 15 metadata elements as touchstone to support simple description and information discovery for all artifact types (“*Dublin Core Metadata Initiative*”, n.d.). It is implemented together with other non-DC elements to enable further description of the artifacts or exhibition. As such, each artifact type will have its own set of metadata. The exhibition metadata together with the photograph and
text metadata, which forms the basic artifacts used in virtual exhibitions, are described briefly as follows.

**Exhibition Metadata**

Exhibition Metadata contains information about the virtual exhibit content. The DC metadata elements are used to provide general descriptions about the exhibit. An exhibition is sub-divided into parts, each with its own part ID and title. In each exhibition part, there should be at least one page element. Each page element has its own page ID, title and defines the different photograph and text artifact presented in the exhibition page as well as the navigational links that links each exhibition page to another. Since layering of information is implemented, photograph accession ID and image ID is used to uniquely identify a specific photograph version while text ID and content ID is used to uniquely identify a specific text version. The photograph element also contains the height and width attribute in pixels which can be optionally defined by the author to determine the image display size in the HTML page. Figure 3 shows the exhibition metadata document type definition (DTD) and Figure 4 shows an example of a typical exhibition’s metadata.

**Artifact Metadata**

The photograph metadata set is used to describe a photograph artifact as shown in Figure 5. The Accession ID is used to uniquely identify a photograph artifact while the location element is used to define the repository directory where all the images are stored. The image element can be defined more than once to cater to the different available versions of the same original photograph. These versions may contain resized, enhanced, or digitally-manipulated variations (e.g. colour, addition of borders) of the original photograph. The same approach has been adopted for the textual artifact metadata to incorporate the standard DC elements plus text specific elements. For example, the “content_version” element is used the support layering of information through different descriptive layers of a textual artifact (i.e. different levels of text description ranging from summarised abstract information to detailed information).
Figure 3. Exhibition metadata document type definition
Figure 4. Example of exhibition metadata XML document
System Implementation

NAS-VES is implemented using Java and XML related technologies. It runs on the Windows Server 2000 operating system, Java Web Server and Software AG’s Tamino XML Server (“Tamino XML Server”, 2002). The Tamino XML Server provides the storage, indexing, and retrieval engine for the different artifact metadata stored in the repositories. The Java programming language is used to build the different classes and servlets while HTML, Java Server Page and Java Script are used to build the Web pages. The actual artifacts and application repositories are external to the server with each XML document uniquely identified and referenced for use by all applications of the digital archive. The server provides the indexing mechanism for all metadata in these XML documents in order to support generic search and retrieval operations.

HTML forms are used by exhibition authors to define the metadata information and exhibition authoring. In the current version of VES, the Metadata Creation Tool uses different interfaces to support the metadata definition of the photograph artifact, text artifact and virtual exhibition. Figure 6 shows the exhibition definition interface. The
top half is used to define the required information while the lower half shows the created XML content. This authoring concept is common throughout all VES interfaces. A similar set of interfaces are used to define the photograph and text metadata (not shown). Through these interfaces, the entire exhibition is defined in an XML document and subsequently indexed in the XML server.

The Virtual Exhibition Authoring Tool defines the layout of the artifacts and creates the “look-and-feel” of the exhibition. Each artifact is positioned in the HTML page using a grid-pattern layout. The grid-pattern is basically a table comprising a number of columns and rows with each element in the table providing a slot to house an artifact display. The final step is to define the CSS and XSL templates (style sheets) to be applied to the page content from a list of available templates. Varied exhibition presentation designs are achieved by using different templates. Additional customised templates may be created, stored in the templates directory and automatically made available for future use. Figure 7 shows an example of the layout out of artifacts and an example of the final HTML page with an applied template. Figure 8 shows an example of an online gallery where all the photographs used in specific exhibitions can be collated together and displayed in a page using their thumbnails so that visitors can download or obtained printed postcard version from them.
Figure 7. Exhibition authoring tool and CSS/XSL style sheet application to achieve final HTML exhibition page

Figure 8. Online gallery page
The final exhibition emerging from VES can be further enhanced by the exhibition author, if necessary, by using traditional Web-page authoring tools such as MS FrontPage (“Microsoft Frontpage”, 2002) and Macromedia’s Dreamweaver (“Macromedia Dreamweaver MX”, 2002). Such enhancements are generally intended to be minor and cosmetic in nature, although it can also be substantial involving major layout and other changes. However, the contents of the pages must be left intact to ensure a one-to-one relationship between these and the information stored in the digital archive. Additional text added in this additional authoring step will not be indexed by VES and could result in inaccurate search outputs if the corresponding content integrity is violated.

When visitors are interacting with the exhibition pages, they have the option to click on the different artifact to display the full record of the detailed metadata and other associated information of the artifact on a new window. Figure 9 shows an example of a selected photograph artifact record. A similar display is shown for the text artifact to reveal the text metadata and different layers of text descriptions (not shown).

Figure 9. Photograph artifact detailed HTML page
The search and retrieval in the virtual exhibitions is achieved through the *Virtual Exhibition Search Tool*. Searches are supported for exhibition, photograph or text artifact individually or across exhibitions through both field and free text search. Field search is intended to support search of known metadata information that is associated with individual exhibition artifacts or exhibition data. Each field corresponds to the XML element of the artifact’s metadata. Free text search is provided to allow natural language searches across all available fields and content of the virtual exhibitions. Boolean operators are supported both in the field search and free text search. The Boolean operators are used for multiple search fields and multiple search terms. Figure 10 shows an example of a search output with the results containing links to exhibition and artifact details pages.

![Figure 10. Typical search result](image-url)
Discussion

VES adopted the *reference and reuse* model in preference over the traditional *copy and use* model for creating virtual exhibitions. In the traditional model, each exhibition obtains a copy of all the artifacts used in that particular exhibition implying that when an artifact is used in multiple exhibitions, multiple copies of it are made and presumably stored in different directories. This makes the process of updating difficult and storage inefficient. In contrast, the *reference and reuse* model allows one unique copy of each digital artifact to be created, stored in the digital archive and referenced by different exhibitions and other applications. This is achieved VES as a result of the digital archive architecture.

Although the process of creating a virtual exhibition using VES appears cumbersome initially, the advantages of the maintenance of the integrity and consistency of artifacts, separation of content with layout in the authoring, and support for an integrated search across repositories far outweighs the initial effort in adapting to the workflow of the system to create exhibitions.

The clear separation of content and formatting information supported in this system fosters the creation of multiple exhibition versions. Multiple versions to support different user interests and user needs may be implemented in three ways. The first is to have the same exhibition content but different formatting templates (i.e. different look and feel). The second is to adapt and satisfy the different information needs of different types of users by drawing upon the appropriate content of the artifact for display (e.g. a child’s exhibit to a serious researcher’s exhibit). The third is to vary both the content and formatting templates to achieve other customised exhibits. The use of CSS and XSL templates achieves and simplifies this process. CSS templates are employed to generate exhibitions with different designs while XSL templates are used to generate different content layout and selection of appropriate levels of content. Transforming XML documents containing the exhibition content into HTML documents through the XSL templates has also shown to ensure the longevity of the exhibition, making it accessible for future generations (Lie & Saarela, 1999). These aforementioned features of VES make it possible to support creation of different virtual exhibitions to support different user needs through one single system, thereby making it flexible, useful and unique in comparison to existing virtual exhibition systems.

VES utilised the standard DC metadata elements to define the basic descriptors for different artifact types and exhibitions. Additional non-DC metadata elements were created to provide a richer and more complete metadata description as necessary. This lends support to interoperability and direct exchange of information in part as a result of using DC. Some form of mapping may also be applied for non standard elements during information exchange to achieve a higher degree of interoperability across different systems.

The search in VES is enhanced by the inclusion of metadata for all artifacts and exhibition data so that it is not only confined to the visible displayed content of the
exhibition. The implementation of metadata therefore increases the number of information access points which is further enhanced by Boolean field searches. As such, we would expect visitors to achieve a higher level of data recall in their searches although we have not done any experiments to verify this claim. VES also adds value to the exhibition by the generation of detailed artifact information that can be displayed by visitors on demand. Such detailed information offer visitors a richer educational experience while viewing the exhibits. The concept of the layering of information can be extended to incorporate even more successive layers of information content to even a mini-package of value-added features associated with an artifact. For instance, links to related resources or other artifact types are examples of such features. A visitor who is viewing a static photograph may be able to activate a video or sound clip that is associated with the photograph, or have access to a series of related photographs of the same subject. The same digital archive architecture can support such forms of extensions and enhancements in future.

The VES authoring system provides a GUI tool to simplify the creation of virtual exhibitions so that exhibition authors with little or no knowledge of web development technology are still able to create simple virtual exhibitions. However, it is acknowledged that the current grid layout method for authoring is limited and lacks flexibility to cater to a range of design needs. As such, it is envisaged that future versions of VES will support other authoring paradigms such as the direct manipulation graphical user interface to allow exhibition authors to have additional flexibility to interface with the digital archive to support the authoring process and to have an alternative means to define the exhibition metadata directly through this interface. Additionally, some provision to allow exhibition authors to manipulate and artistically enhance individual artifacts during the exhibition creation will also be useful.

Although the created HTML exhibitions through VES lack the interactivity of multimedia versions, the resulting exhibition can be enhanced to support such interactivity by adding scripts into the XSL templates. However, it must be cautioned that the addition of scripts make the exhibition device-dependent since different browsers support different proprietary scripting languages.

Although VES was created to support the digital archive of NAS, we believe that this system is portable and may be implemented by other organizations to create and maintain their virtual exhibitions since all virtual exhibitions use similar artifacts. Furthermore, the properties of these artifacts are completely captured in its metadata and these properties are universal in nature in which an organization can readily adopt and use. Different templates may be created to suit the “look and feel” that particular organisation wants to achieve for their virtual exhibition.

**Conclusions**

A virtual exhibition system, VES, has been successfully designed and implemented to support the creation of virtual exhibitions containing artifacts from the NAS digital archive. The XML-based digital archive architecture provides a suitable platform for
many other archive and heritage related applications to be created to serve the many varied user groups it services. This initial version of VES provides a methodology and process for creating simple virtual exhibits that can be further enhanced by exhibition authors as necessary. The strength in VES lies in its simplicity, adoption of the reference and reuse model, support of interoperability through the use of DC metadata as the foundation to describe its contents and artifacts, search capability of the system and value-adding feature of providing additional content on-demand. VES will go through a series of evaluation and usability tests to surface areas for refinement and improvements. Future work envisaged in VES include the provision of alternative authoring paradigms, inclusion of additional artifact types, facilitating enhanced interaction and interactivity in the exhibitions, and providing a digital work environment for visitors in interacting with the exhibitions. The latter would allow virtual exhibitions to be value-added and utilized more productively to accomplish users’ tasks instead of merely providing a platform for the dissemination of information.

References


Lim, J.C. (2002). A virtual exhibition system for the National Archives of Singapore. Master of Science (Information Studies) dissertation, School of Communication & Information, Nanyang Technological University, Singapore.


