

A Paper submitted for the Digital Library for the Maghreb Workshop
Held in Rabat, Morocco, January 25-27, 2007

Digital Library of the Maghreb: A Perspective

Cinda Ann May, Cunningham Memorial Library, Indiana State University.
Paul Asay, Cunningham Memorial library, Indiana State University.

Today, the countries of the Maghreb region, Mauritania, Morocco, Algeria, Tunisia, and Libya have the opportunity to create a fundamental building block of a rich cultural heritage. The idea of a Maghreb Digital Library will positively be a catalyst to research, teaching, and learning in the areas of cultural heritage resources held by institutions, and individuals around the world.

This paper is divided into two parts:

1. An example of a digital library model implemented at Indiana State University.
2. Maghreb Digital Library in perspective

Part 1: Indiana State University Project
“The Wabash Valley Visions & Voices”

The Wabash Valley Visions & Voices Digital Memory Project was initiated in May 2004. Area organizations came together to form a partnership for the specific purpose of creating a digital repository of local history and culture. The mission of the project is dedicated to the documentation and the preservation of the history and cultural heritage of west central Indiana.

The project is a collaborative effort by the region’s libraries, cultural organizations, community groups and area residents. The digital assets created by the participants are made freely available through a searchable database located on the project’s website at <http://visions.indstate.edu>. The database contains photographs and moving pictures, sound files, artifacts, manuscripts, oral histories, yearbooks, and printed texts, which are both historic and recent in origin. At present, eighteen organizations across five Indiana counties contribute material to the database, which contains approximately 32,000 items. The project also serves as a “Resource Partner” for the Indiana Humanities Council’s smartDESKTOP Initiative, which furnishes digital assets in support of K-12 instruction. In the future the contents of the “Visions Project” will also be available through “Indiana Memory,” a statewide digital library currently under construction.

The “Visions Project” is based at Indiana State University (ISU) Library, which provides the collection management software (CONTENTdm), the server space for storage, and specialized staff members, who assist in the digitization of material and metadata creation. Currently, the ISU Library project staff consists of the Coordinator of Library Digital Initiatives, who serves as the project manager, a Metadata/Digital Initiatives Librarian, a Metadata Specialist and student employees. The ISU Library maintains a digitization lab with four workstations and scanners, including a large format scanner, as

A Paper submitted for the Digital Library for the Maghreb Workshop

Held in Rabat, Morocco, January 25-27, 2007

well as the hardware and software required to produce streaming media files. The lab and all of the equipment are available for use by the content providers. The initial CONTENTdm training was outsourced, but all subsequent training is conducted by ISU Library personnel.

The ISU Library purchased a CONTENTdm license at the 32,000 objects level, which it subsequently upgraded to “unlimited” status. The CONTENTdm collection management software resides on an ISU Library server. Through the software’s Acquisition Station participants import digital files and create searchable metadata according to established standards. A partner may choose to have one or more acquisition stations installed on site for its use and upload the digital assets it creates to the ISU Library server. In some instances, members of the ISU project staff perform the digitization and/or metadata creation tasks for partners. This may require the transportation of source material from a participant’s site to the ISU Library. In most cases, the Coordinator handles this task. All original material is stored in a secure place during the digitization process and is returned to the owner in a timely fashion. Each contributor determines the content of its collection and is responsible for copyright clearance. The Coordinator of Library Digital Initiatives maintains a central file for copyright permissions.

A CONTENTdm license holder is required to pay a yearly maintenance fee. The founding partners of Wabash Valley Visions & Voices agreed to divide the annual cost according to a sliding scale based upon each participating institution’s annual operating budget. This partnership fee schedule is part of a “Memorandum of Understanding” established by the founding partners that sets forth the responsibilities and expectations of all participants. A partner uploads its digital assets into a discrete collection within Wabash Valley Visions & Voices, and is represented by a flash page that provides institutional contact information, a collection description with imbedded links to topics of interest, a collection specific browse and search feature, and a link to the organization’s homepage.

The ISU Coordinator of Library Digital Initiatives represents the “Visions Project” on the local, state, and national levels. The coordinator is responsible for the development and maintenance of the project’s collaborative regional network, the identification of funding sources and composition of grant proposals, project promotion and special events. In addition, the coordinator oversees the ISU project staff, communicates with partners and facilitates the digitization and metadata creation process. The founding partners meet several times a year to discuss issues that affect the project and for professional development opportunities. Publicity about the Visions Project includes all contributors regardless of the focus as Wabash Valley Visions & Voices is an endeavor that each institution has made a commitment to support.

Grant funding to develop digital assets may be sought by a single organization or a group of institutions working together to document a specific topic. All individual revenue streams are carefully preserved and questions concerning items within a collection are referred to the appropriate repository. All metadata records contain contact information

A Paper submitted for the Digital Library for the Maghreb Workshop

Held in Rabat, Morocco, January 25-27, 2007

for the repository as well as a copyright notice. In addition, community groups with small collections and area residents may participate on a no-fee basis by contributing material to a county specific community collection. For example, John Smith of Parke County works with a member of the Visions Project staff who digitizes the Smith Family photographs and creates the metadata based on information provided by Mr. Smith. The digital assets thus created become part of the Parke County Community Collection that contains documents and photographs submitted by other contributors from the county. The repository element of the metadata states: "From the private collection of John Smith, Parke County, Indiana." To develop these county community collections Wabash Valley Visions & Voices periodically invites the public to bring material to a specified place, usually a public library, on a designated day to be scanned and recorded for the project. This enables project staff to capture the information to create the digital asset while the item remains in the possession of the individual. These "Community History" days are a means to publicize the project and promote public involvement.

The Wabash Valley Visions & Voices Digital Memory Project homepage presents the visitor with a variety of options for searching and acquiring information. A "Current Item of Interest" feature highlights new collections or special topics and is changed approximately every month. In this manner the site remains fresh and invites return visits. The "About Us" section includes project documentation, a copyright statement, photographs and text describing events, a tutorial and links to other repositories. The "Collections" section provides links to each partner's and county community collection that may be browsed or searched from a flash page. The "Subjects" section offers twenty subject categories that retrieve digital assets across all collections. This subject browse feature is based on categories selected by the developers of "Indiana Memory" and is supplemented by topics of local interest. Past "Current Items of Interest" are relocated to the Subjects page when replaced by new material. As the Visions Project is a resource provider for the Indiana Humanities Council a link to the smartDESKTOP Initiative is furnished to highlight that partnership.

Public response to the Wabash Valley Visions & Voices Digital Memory Project is overwhelmingly positive. The Indiana Library Federation presented its 2006 Collaboration Award to the project, which is considered a model for other collaborative efforts. The project respects the individual identities of the contributors while it fosters a group persona. Area residents are encouraged to contribute photographs and documents as a means to nurture civic pride and a sense of community. Future plans for the Wabash Valley Visions & Voices Digital Memory Project include the expansion into additional Indiana counties, the creation of contextual material to enhance classroom use of the digital assets, and the development of usability testing. Wabash Valley Visions & Voices will continue to evolve and embrace new ideas and technology as it documents and preserves the history and culture heritage of the region.

A Paper submitted for the Digital Library for the Maghreb Workshop

Held in Rabat, Morocco, January 25-27, 2007

Part 2: Maghreb Digital Library in Perspective

A great opportunity is available for the Countries of the Maghreb to establish working partnerships and networking to digitize their rich cultural heritage, preserve it, and teach it to future generations. In this regard, it will be important for all the countries to engage in discussions pertaining to the issues of designing and implementing a Digital Library. The following are some factors that I think should be addressed:

1. Define the role of each country in creating a contributing institution (University in preference) that mounts its content locally (collecting information, storing it, and making it accessible electronically).
2. Provide each contributing institution with the necessary infrastructure, equipment, and expertise for metadata harvesting and scanning to insure reliable repository for digital materials.
3. Create a leading institution in one of the countries that will host the Maghreb Digital Library.
4. Itemize and align the roles of the leading institution with best practices.
5. Create very high speed networks to enable all the institutions to work collaboratively.

Other Factors to consider:

1. Expertise and technical support
2. Infrastructure and space allocations
3. Administrative and management support
4. Financial support (Funding, revenue generated streams)
5. Staff and managers training and development
6. Languages

Possible mission statement:

Maghreb Digital Library is dedicated to collection, documentation, and preservation of the digitized cultural heritage of the Maghreb countries of Mauritania, Morocco, Algeria, Tunisia, and Libya. As a collaborative endeavor among all the countries in the Maghreb, the Library provides free access to individuals and institutions around the world via the Internet.

Possible Technical background

A variety of Digital Library Management systems are available to operate a Digital Library. Very brief summaries are presented on two such systems currently available:

DSpace:

DSpace was developed through a collaboration of MIT Libraries and Hewlett-Packard Labs. DSpace is freely available as open source software. Individual items within DSpace are grouped into communities and collections of items. Descriptive metadata describes each item within these communities and collections. DSpace uses only a web browser to insert items into a collection. Searching the system is fairly straight forward via a web browser.

A Paper submitted for the Digital Library for the Maghreb Workshop

Held in Rabat, Morocco, January 25-27, 2007

Drawbacks to DSpace relate more to it's complexity as an open source tool. Most of the tools require some knowledge of LINUX/UNIX operating environment and tools. See <http://www.dspace.org> for more detailed information.

ContentDM:

ContentDM as opposed to DSpace is a for profit based product that requires an annual subscription fee. The fee is based on the number of items held within the collections. ContentDM stores individual items within collections. Items are added to the collection via a Windows based software tool they call an Acquisitions Station. Each item is given descriptive metadata, and then submitted to the server for approval to be added to a specific collection.

The advantage of ContentDM is a dedicated development and support staff. It tools are well developed and updated on a regular basis. The disadvantage over DSpace is the cost. See <http://www.contentdm.com> for more detailed information.

Server:

The server consists of a hardware platform running a UNIX or LINUX operating system. Some Digital Libraries can also utilize Windows Server 2003 in place of UNIX or LINUX.

The server hardware itself is usually based on fairly inexpensive x86 based processors. Multiple network connections, power supplies, external SCSI interfaces and internal storage drives would need to be included within the server.

External storage disk arrays connected to the server via SCSI connections are typically configured to hold the digital media itself. Digital Libraries over time usually contain a vast amount of storage needed for constantly growing and diverse digital collections.

A typical server configuration as an example that meets the above requirements would be: Sun V40Z Server (Config 1) with a based price of \$7,025.00 US

Storage Array:

Storage arrays are used to hold the digital information itself. These arrays contain multiple disk drives and connect to the server via a host buss adapter and a SCSI cable. The information is maintained on multiple disks (RAID array) so that if any one disk fails it will available immediately from a spare drive within the array. The failing disk can them be replaced at any time. The costs of these arrays vary greatly based on capacity, fail over ability and speed.

A typical storage array configuration that meets the requirements for a medium sized digital library would be:

Sun 3320 SCSI Array with 1.7 terabytes of storage, base price \$27,335.00

Sun X9265A Storage Controller (to connect server to array) base price \$650.00 US

Backup:

In case of potential disaster recovery should the server site be subject to fire or water damage, a backup should be performed on the server. For best practice at least one full

A Paper submitted for the Digital Library for the Maghreb Workshop

Held in Rabat, Morocco, January 25-27, 2007

backup tape should be stored off site. Backup tape drives and the tapes themselves come in a variety of capacities and prices.

A typical backup tape drive to support a medium size digital library would be:
Exabyte 7192 LTO Autoloader LTO-2 rack mount loader, base price \$4,725.00

Server Room Itself:

Servers are typically housed in rooms that provide clean power and controlled temperatures. Most server models are constructed to be stored in a standard sized rack mounting enclosure.

Client Workstations:

Digital Library software on individual workstations is responsible for acquiring digital resources, the information about the resource, and then transferring it to the server. A personal computer based on the Windows operating system is used as the client. A high speed network connection to the Internet via TCP/IP would be required on each of these client workstations. Any number of client workstations can be used to add digital content to the server.

Peripherals to Digitize Documents and Photographs:

Client workstation often requires a digital scanner attached to scan photographs and documents into digital form to be then uploaded to the server. A variety of scanners are available for different size materials with a wide variety of prices.

A typical scanner used for digital library projects would be:
Microtek 9700xl large format scanner base price \$999.99 US

In addition a variety of tools may also be required to capture and edit audio and video sources to be included within collections.

Specific Challenges:

- Both digital library management systems support Latin based character sets, but do not directly support the Arabic language.
- Initial setup and the inevitable software upgrades needed to support the server based software require some technical assistant with a person familiar with LINUX/UNIX based tools.
- Depending on the scope and quantity of digital resources added to the various collections may require additions of storage capacity along with the potential for multiple server environments.
- Knowledge of the various digital formats and the standards used in the creation of digital resources is a highly recommended advanced knowledge before embarking in creating collections.