GETTING STARTED WITH DIGITAL COMMONWEALTH

Digital Commonwealth (www.digitalcommonwealth.org) is a Web portal and fee-based repository service for online cultural heritage materials held by Massachusetts libraries, museums, historical societies, and archives. The portal facilitates the searching and browsing of member institutions' digital assets, allowing the discovery of all kinds of digital resources including manuscripts, images, historical documents, and sound recordings. The repository provides a storage service for institutions that choose not to host their own digital assets. The metadata stored in the repository will be indexed by the portal.

Building a digital collection to be accessed via the portal requires thoughtful planning. What follows is a list of steps to be considered when planning a digital project.

1. Selecting Items or Collections to Digitize

Why digitize? There are numerous reasons:

- Meeting user expectations
- Providing enhanced access to institutional holdings
- Increasing exposure of your institution and its collections
- Minimizing handling of original source materials
- Contributing content to collaborative digital projects

Please note: digitization is an *access* tool, not a preservation method. (For more information about preservation and sustainability issues relating to digital assets, see section 7 "Preserving Digital Assets" below.)

Some questions to address:

- Who owns the copyright to the material?
- Is the material in demand? Will digitization likely increase its use?
- Can the materials be safely digitized? (Fragile? Oversized? Brittle? In a problematic format?)
- Will digitization provide access to materials that might otherwise be limited due to preservation concerns?

Once the institution has selected the items or collections to digitize, document the decision-making process — the rationale for what was chosen, staff involved in the process, the date -- as this may inform future selection strategies.

2. Organizing and Describing the Selected Item or Collection(s)

Organizing materials according to their provenance and original order protects their context and achieves physical or intellectual control over them. Describing the item or collection(s) facilitates access and creates a record of the organized collection(s) and its components. These descriptive tools (finding aids, inventory lists, catalog records) often include container lists and notes about the title, size, dates, and scope and content of the collection.

A well-organized and described collection enhances access to the materials, increases its value, and facilitates the selection process for digitization.

3. Creating Digital Assets

Selecting an appropriate approach to digitization is at least as important as selecting the items or collections to digitize. Although scanners are easily available, creating digital images on a readily available desktop scanner is not the only option; in fact, using a reputable contractor, rather than digitizing in-house, might be a better choice. To create digital images of maximum utility to the user requires knowledge of image specifications, scanner hardware capabilities, software configurations, image quality control issues, and the ability to create appropriate and complete metadata. It is possible to create what appears to be an adequate digital image, but sometimes even large digital images (with what appear to be appropriate pixel dimensions) might not be usable if they do not have good tonal range and color balance. Either way, digitizing in-house or working with a vendor, establishing digital image specifications beforehand ensures quality and functionality of the end product.

Please note: When the Digital Commonwealth's repository service policy has been finalized, the recommended specifications will be stated here.

Specifications used by the Northeast MA Regional Library System for a digital repository project:

Scan Size (ratio) of Original Item: 1:1 ratio
The scanned image should not be manipulated to increase or decrease its output size.

Resolution Bit-Depth
Color Photographs: 24 bit
Black & White Photographs: 8 bit gray scale
Black & White Text: 1 bit

Resolution DPI
Digital Master: 600dpi
Access Image (for printing or detailed web viewing view): 300dpi
Web Viewing: 72dpi

Image File Formats
Digital Master: Uncompressed TIFF
Web Images: JPEG
Access Image (for printing or detailed web viewing): JPEG

For another example of digitization specifications, visit C/W MARS' Digital Treasures:
http://dlib.cwmars.org/cdm4/images/cwmars_benchmarking.pdf

Digitizing In-House or Working with a Vendor--Indispensable Components of any Digital Project

- **Infrastructure**: ensure administrative capacity, including adequate funding, staff expertise, and technological capacity
- **Staff time**: budget sufficient staff time to administer the project (many institutions find that they have not considered this seriously enough before beginning a project)
- **Materials handling**: attend to security, shipping, insurance, climate control, tracking, experience working with fragile/special materials
- **Metadata**: determine type and level of required metadata and how they will be created and stored (for more information, please see section 4. "Creating Metadata" below)
- **Quality control**: review the digital images and metadata to confirm that they meet technical specifications
- **Sustainability**: maintain long-term storage of digital asset, data migration, file integrity, interoperability, functionality of the digital asset, hardware and software obsolescence (it happens faster than you think!), and adequate on-going funding and support

Digitizing In-house or Working with a Vendor --Things to Consider When Weighing Both Options:

	Working with a vendor	Digitizing in-house
Project management	Locate a suitable vendor and develop a project agreement (scope; cost; technical specifications; shipping, storage, security, and handling requirements; and project schedule.) The vendor is responsible for meeting all contractual obligations.	Determine the project schedule and staffing, including project supervision. Internal staff must coordinate and complete all tasks on time and meet required technical specifications.
Funding	Have sufficient funds to pay the vendor.	Have sufficient and reliable funding over time to cover staff, equipment, maintenance, and supplies costs.
Staff expertise	Maintain ongoing communication with the vendor to ensure compliance with technical specifications. The vendor is responsible for having knowledgeable staff.	Have production staff who can create digital images according to technical specifications, and supervisory staff who can manage digitization projects.
Technological capacity	Vendor maintains up-to-date hardware, software, and network infrastructure needed to meet all contractual obligations. (Note: your institution needs a mechanism to ensure quality control. This may require some equipment for staff to review images prior to paying the vendor.)	Buy, configure and maintain hardware, software and network infrastructure needed to create all necessary digital products. Responsible for quality control, including review of digital images.
Materials handling	If imaging is done offsite (sometimes vendors work onsite), the vendor is contractually responsible for secure and appropriate storage and handling of materials, and for packaging and return shipping.	Handle materials appropriately during digitization.
Hardware/software obsolescence	Vendor ensures its hardware and software is up to date.	Maintain up-to-date hardware and software.
Digital storage / sustainability / preservation plan	Vendor is responsible for adequately storing the digital files during the time of the contract. Institution must ensure file maintenance, data migration, sustained access to, and preservation of digital assets after completion of the project.	Responsible for a sufficient backup system for all digital files during image production and beyond. In addition to backup systems, create a digital preservation plan that should include data migration, file integrity checking, continued functionality, refreshing and reformatting the data.
Metadata	Vendor will provide metadata as per the contract.	Responsible for descriptive, administrative, and structural metadata.

4. Creating Metadata

After they are created, digital assets are stored on a host server, either within the home institution or within a repository service. To be accessible, these digital assets require metadata, which is information about or pertaining to the digital assets (object type, size, date, creator, subject, provenance, etc.). Metadata facilitates data management, enables access, and enhances sustainability. The more complete the metadata, the easier it is to manage the digital assets.

Metadata typically fall into several categories, including ***administrative***, ***descriptive***, ***preservation*** and ***structural***. Dublin Core (www.dublincore.org) is an important descriptive metadata standard that allows for the discovery of digital assets. The full Dublin Core standard includes fifteen elements or fields which are listed below. Recommended best practice is to use a content standard, thesaurus, or encoding scheme for each element. The elements in bold are particularly useful for the retrieval of digital assets by the Digital Commonwealth.

- **Title (required by)**
- **Creator**
- **Subject**
- **Description**
- Publisher
- Contributor
- **Date**
- Type
- Format
- **Identifier (required by)**
- Source
- Language
- Relation
- Coverage
- Rights

For the portal to provide access to digital assets from multiple repositories, it must harvest (or gather) Dublin Core metadata related to those assets in a predictable and structured manner. An international standard known as the Open Archives Initiative (OAI) (www.openarchives.org) uses Dublin Core elements and defines the requirements for making descriptive metadata harvestable. The portal requires compliance with the OAI standard.

5. Making Digital Assets Available on the Web

Making digital assets available on the Web can be accomplished in more than one way:

- from an institution's web site that is configured to manage digital assets and metadata;
- if the institution does not have a web site, the digital assets and metadata will need to be housed in a digital repository that can be harvested by the Digital Commonwealth portal.

Options include:

- Subscribe to the Digital Commonwealth's repository and make your institution's assets and metadata available through the Digital Commonwealth portal.
- Store the digital assets and metadata in a shared repository hosted by a Digital Commonwealth member (e.g. C/W MARS, NOBLE). Options for participation in a shared repository will vary depending upon your institution's consortial memberships.

In either case, costs will be associated with storage of the records.

Please be aware that if the project is outsourced and the vendor is able to provide an affordable repository for your digital assets, they may not be searchable in the Digital Commonwealth portal unless the vendor is a member.

6. Contributing Metadata to Digital Commonwealth

Members of the Digital Commonwealth contribute OAI-harvestable metadata to the **portal** in a variety of ways depending on where their digital assets are stored.

If the digital assets are stored in:

- The **Digital Commonwealth repository:** the corresponding metadata are systematically included in the Digital Commonwealth portal
- A **member repository** (e.g. C/W MARS or NOBLE): the corresponding metadata are harvested by the Digital Commonwealth portal
- A **member institution's server and is accessible through its web site:** if the institution creates OAI-harvestable metadata, it can be harvested by the Digital Commonwealth portal. The member institution must send the baseURL (the web address to be used by the harvester if the institution is creating OAI metadata on an ongoing basis) or the Static Repository URL (the web address of a specific completed set of OAI-harvestable metadata) to the Digital Commonwealth portal administrator.
- A **vendor's repository:** if a Digital Commonwealth member stores its digital assets in a vendor's repository and if that repository produces unique URLs to the digital assets and creates OAI-harvestable metadata, it can be harvested by the Digital Commonwealth portal. As in #3 above, the vendor must send the baseURL (the web address to be used by the harvester if the institution is creating OAI metadata on an ongoing basis), or the Static Repository URL (the web address of a specific completed set of OAI-harvestable metadata) to the Digital Commonwealth portal administrator.

7. Preserving Digital Assets

Digitization is a way to provide access to cultural materials that are fragile or rare by the creation of a digital surrogate; but digital resources must also be preserved. Some of the preservation issues to be aware of in the long-term sustainability of digital objects are:

- Poor quality digital capture
- Inadequate preservation metadata
- Bit degradation of digital files
- Obsolescence of hardware and/or software platforms
- Hardware and/or software failure
- Insufficient infrastructure to support long-term storage
- Insufficient long-term commitment of human, technological, and financial resources
- Use of unsupported, proprietary file formats
- Deficient back-up of data
- Poorly implemented migration strategies
- Human interference – accidental or deliberate

Check with the repository service provider regarding their policies on preservation of digital assets. General guidance on preservation issues, including resource links, is provided at www.digitalcommonwealth.org/docs/preservation/ .

8. Prepare for Popularity!

Making collections searchable and available on the Web may generate new attention for an institution. A whole new audience can discover collections, and some will want to visit in person. Research shows that if you build it, they will come!