

Exercise 1: Developing Achievable Goals

Our archival repository is developing a grant proposal so that we can digitize the papers of four families: the Brown, Black, White and Orange families. The papers date from 1832 to 1898. They include diaries, letters, photographs and legal documents. There are approximately 56,000 pages. The collections are fully processed and have been open to researchers for several years. Almost all of the materials are hand-written. Some of the diaries have been transcribed because they are hard to read.

There are paper finding aids for all four collections. The finding aid for the Orange family papers is online in HTML, and there is a catalog record for the collection. We hope that this project will result in online finding aids and catalog records for the other three collections. The finding aids are very detailed and include item-level descriptions for the large majority of the contents. Descriptions of the diaries have been entered into a simple database.

Our target audience includes scholars, students, genealogists, and the general public, but we really are focusing on the first two groups. Scholars and graduate students have used the collections for years, but it would be good to reach out to undergraduate students and even K-12 students.

We have most resources to complete the project in-house (expertise, hardware, space), but lack sufficient funds to hire the staff needed to complete the work. We will make all digital objects freely available online using our existing digital collections system. This will be an 18-month project.

Assignment:

1. As a group, create at least 4 goals for this project.
 2. List specific activities that will accomplish each goal.
-

Exercise 2: Metadata Mapping Exercise

Assignment: As a group, create a metadata map that takes source metadata that is relatively rich but non-standard and converts it to simple Dublin Core, which will be used by all partners in a collaborative digital project.

Questions to ask: Which fields map directly? Will data need to be “massaged” in order to work in Dublin Core? Will we capture all data or will some data be lost?

Source metadata: A diary from the Orange Family Papers described in an MS Excel spreadsheet created by a scholar. The name of each field in the database appears in all caps in the left column. The contents are the fields are in the right column.

DIARY NUMBER	12
DIARY DATES	1865-1869, plus a few entries from 1873
DIARY TITLE	On cover: Journal of H.M. Orange. On first page: H.M. Orange’s Daily Diary
DIARY AUTHOR(S)	H.M. Orange
DIARY PAGE COUNT	96
TRANSCRIPTION	No
DIARY SUBJECTS	Civil War, Plantation, Politics, Lincoln, weather, land
LOCATIONS	Jacksonville, St. Augustine, Tallahassee
DIARY SUMMARY	H.M. Orange primarily discusses family matters, management of the plantation, and personal business (land transactions). There is discussion of the war, Lincoln, slaves, and national politics. Includes daily weather summary.

Target metadata: Simple Dublin Core, which includes the following elements:

Title - The name given to the resource. If in doubt about what constitutes the title, repeat the Title element and include the variants in second and subsequent Title elements.

Example: Title="The Sound of Music"

Creator - An entity primarily responsible for making the content of the resource. Personal names should be listed surname or family name first, followed by forename or given name. When in doubt, give the name as it appears, and do not invert.

Example: Creator="Shakespeare, William"

Subject - The topic of the content of the resource. Typically, a Subject will be expressed as keywords or key phrases or classification codes that describe the topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme (e.g., LC Subject Headings). If multiple vocabulary terms or keywords are used, either separate terms with semi-colons or use multiple Subject elements.

Example: Subject="Aircraft leasing and renting"

Description - An account of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Publisher - The entity responsible for making the resource available. Examples of a Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.

Example: Publisher="University of South Where"

Contributor - An entity responsible for making contributions to the content of the resource. Examples of a Contributor include a person, an organization or a service. Typically, the name of a Contributor should be used to indicate the entity. The same general guidelines for using names of persons or organizations as Creators apply here.

Date - A date associated with an event in the life cycle of the resource. Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value follows the YYYY-MM-DD format. If the full date is unknown, month and year (YYYY-MM) or just year (YYYY) may be used. Many other schemes are possible, but if used, they may not be easily interpreted by users or software.

Examples: Date="1998-02-16"; Date="1998-02"; Date="1998"

Type - The nature or genre of the content of the resource. Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary.

Examples: Type="Image"; Type="Text"; Type="Diary"

Format - The physical or digital manifestation of the resource. Typically, Format may include the media-type or dimensions of the resource. Examples of dimensions include size and duration.

Information concerning the size of a resource may be included in the content of the Format element if available.

Examples: Format="3x10"; Format="3 pages"

Identifier - An unambiguous reference to the resource within a given context. Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Examples of formal identification systems include the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).

Example: Identifier=" ISBN:0385424728"

Source - A Reference to a resource from which the present resource is derived. The present resource may be derived from the Source resource in whole or part. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system.

Example: Source="Image from page 54 of the 1922 edition of Romeo and Juliet"

Language - A language of the intellectual content of the resource. Either a coded value or text string can be represented here. If the content is in more than one language, the element may be repeated.

Examples: Language="en"; Language="Primarily English, with some abstracts also in French."

Relation – [summary not included here]

Coverage - The extent or scope of the content of the resource. Coverage will typically include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names. Where appropriate, named places or time periods should be used in preference to numeric identifiers such as sets of co-ordinates or date ranges.

Examples: Coverage="1995-1996"; Coverage="Boston, MA"

Rights - Information about rights held in and over the resource. Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.

Example: Rights="Access restricted to family members"

Exercise 3: Evaluation

Using the list of project goals/activities created in Exercise 1, the group will create an evaluation plan (summative or outcome-based) for our project.