SEPTEMBER–DECEMBER 2021

01–05 METADATA DESIGH ACTIVITY

06 METADATA PROCESSING ACTIVITY

EMMA METCALFE HURST
ARST 556L: METADATA
DR. JULIA BALLARD
FINAL PROJECT

METADATA DESIGN & PROCESSING ACTIVITIES

METADATA DESIGN ACTIVITY

In this short (1-2 pgs), ungraded activity, you will create a small or partial metadata schema for an object or collection of your choice. Please submit through this assignment your sketch (format is up to you; table is most common) and your brief notes on the optional follow-up steps.

- **Specify a use case that includes users, objects, and context (e.g., leisure readers looking for novels in a shared library catalogue)**
-
    - User Group 1: Ceramic collectors / historians looking for information (maker, glaze, location of creation, etc.), about ceramics in an open, crowd-sourced database.
    - User Group 2: Ceramicists who want to share information about their ceramics with ceramic collectors / historians in an open, crowd-sourced database.

- **Make a list of attributes relevant to the objects, users' needs, and the context of the collection. Consider the types of metadata we discussed in Week 2 (descriptive, administrative, technical, structural, and use).**
-
    - **See table figure below**

METADATA DESIGN ACTIVITY

- **Sketch out the details of each attribute. Consider the layers of metadata work from Week 3 and list any relevant details for each attribute, such as:**
  - **Should some attributes have specified formats for entry (for consistency or for easy interoperability with overlapping systems)?**
  - **Should the possible values for some attributes be limited selection from a list (e.g., drop-down menus, controlled vocabularies)?**
  - **Should some attributes be broken up into multiple parts (e.g., title/subtitle)? Do some attributes have relationships to each other (e.g., translated by/translated into/translated from/date of translation)?**
  - **Should some attributes be required, recommended, or optional for entry? Which attributes should be repeatable (e.g., author) and which should occur only once per record (e.g., unique id#)?**
  - **Do some attributes require instructions on how to determine the correct value (e.g., if the title on the cover and title page do not match)?**

    - **See table figure below**

- **Depending on your interest, follow up on your schema with one (or all!) of the following steps and focus on adaptation to standards:**

Metadata Design Table : Canadian Ceramics Database

User Group 1: Collectors / Historians

Needs: Discover biographical information about ceramic artists / discover ceramic artists / identify ceramics

Context of collection: crowd-sourced, grassroots, DIY

Attributes:

User Group 2: Ceramicists

Needs: Share biographical information and pottery with collectors / historians / ceramicists

Context of collection: crowd-sourced, grassroots, DIY

Notes:	Metadata type:	Attribute:	Specified format for entry (for consistency or for easy interoperability with overlapping systems)?	Should values be a limited selection from a list (e.g., drop-down menus, controlled vocabularies)?	Values	Should attributes be broken up into multiple parts (e.g., title/subtitle)? Do attributes have relationships to each other (e.g., translated by/translated into/translated from/date of translation)?	Should attributes be required, recommended, or optional for entry?	Which attributes should be repeatable (e.g., author) and which should occur only once per record (e.g., unique id#)?	Do attributes require instructions on how to determine the correct value (e.g., if the title on the cover and title page do not match)?
Info about user:									
	Use	User's name	Yes ; first and last name	No	n/a	n/a	Required	repeatable	self-explanatory
	Use	User's geographic location (country)	Yes ; for consistency	Yes	List of countries	n/a	Recommended	repeatable	self-explanatory
Info about object:									
	Provenance	Date item was added to collection	Yes ; d / m / y	No	n/a	n/a	Required	unique	self-explanatory
May need a map to help users determine value	Descriptive	*Geographic Region / Country* where pot was sourced	Yes ; regions for consistency	Yes	Japan / China / Korea / West Coast Canada / Central Canada / Eastern Canada / United States of America / Central America / South America / Eastern Europe / Western Europe / Central Europe / Africa / Middle East / Meditteranean / Scandinavia	n/a	Required	repeatable	Yes, include a map to inform users what region / country they are situated in
For local context	Descriptive	*City / Town* where pot was sourced	Yes ; for consistency	No	n/a	n/a	Required	repeatable	self-explanatory
	Descriptive	Pot Type	No	Yes	Unomi / Chawan / Lidded jar / Chalice / Jug / Plate / Platter / Bowl / Vase etc.	n/a	Recommended	repeatable	Yes, include description and visual reference for pot types to allow user to select pot type
	Descriptive	Condition	Yes ; for consistency	Yes	Good / Fair / Damaged	n/a	Recommended	repeatable	self-explanatory
	Administrative	Unique ID	Yes ; for consistency	Yes	automatically generated	n/a	Required	unique	n/a
Info about Object (if known):									
	Provenance	Past Owner(s)	Yes ; first and last name	No	n/a	n/a	Optional	repeatable	self-explanatory
	Provenance	Current Owner	Yes ; first and last name	No	n/a	n/a	Optional	repeatable	self-explanatory
	Descriptive	Maker Name (Ceramicist / Studio)	Yes ; first and last name	No	n/a	n/a	Optional	repeatable	Yes, clarify that "Maker" includes both Ceramic artist and / or Studio maker
	Descriptive	Date of creation (pot)	Yes ; year only	No	n/a	n/a	Optional	repeatable	Yes, clarfy both year and era are possible values
	Descriptive	Creator Education (workshops / artist residencies / schools / apprenticeships)	No	No	n/a	n/a	Optional	repeatable	Yes, explain "Education" includes workshops / artist residencies / schools / apprenticeships
	Descriptive	Clay Type	Yes	Yes	Porcelain (highest fired ; vitrified) / Earthenware (low fired ; porous) / Stoneware (high fired ; nonporous)	n/a	Optional	repeatable	Yes, explain clay types with image examples
	Descriptive	Glaze Type(s)	No	No	n/a	n/a	Optional	repeatable	Yes, explain glaze types with image examples
	Descriptive	Construction Technique(s)	Yes	Yes	Handbuilt / Wheel thrown / Slip cast	n/a	Optional	repeatable	Yes, explain construction techniques with image examples
May need a map to help users determine value	Descriptive	Cultural Origins / Influences	No	Yes	Japan / China / Korea / West Coast Canada / Central Canada / Eastern Canada / United States of America / Central America / South America / Eastern Europe / Western Europe / Central Europe / Africa / Middle East / Meditteranean / Scandinavia	n/a	Optional	repeatable	Yes, explain influences
Media									
		Image of pot	yes ; jpg	n/a	n/a	n/a	Required	unique	self-explanatory
		Image of chop mark / signature	yes ; jpg	n/a	n/a	n/a	Required	unique	self-explanatory
Searching Tools									
		Chop Mark / Stamp Shape(s)	yes ; checkboxes	Yes	Shapes / lines, etc.	n/a	Recommended	unique	Yes, explain how to search the database using these elements
		Chop Mark Letter(s)	yes ; checkboxes	Yes	alphabet / reversed alphabet	n/a	Recommended	repeatable	Yes, explain how to search the database using these elements
		Signature Letter(s)	yes ; checkboxes	Yes	alphabet	n/a	Recommended	repeatable	Yes, explain how to search the database using these elements

METADATA DESIGN ACTIVITY

- **Look up metadata standards likely to be relevant to your use case (e.g., RDA for library materials, DACS for archives) and see how well they fit what attributes and instructions you have anticipated. You can also look into their technical layers (data formats like MARC21 or EAD) on how well they enact the kind of relationships you would think you'll need.**

Because of the nature of the ceramic objects / chop marks / signatures / stamps being described, there is a degree of flexibility needed from the metadata standards. It is important to note that local and specialized qualifiers will also be necessary to include. Research shows that Dublin Core and Categories for the Description of Works of Art (Getty Trust) are two possible metadata standard options. As a universal minimum standard to describe physical and digital objects, Dublin Core Metadata Element Set would work well for this collection. In addition to the 15 "core" metadata elements, Dublin Core includes additional metadata elements to provide for greater specificity and granularity which would allow for more customization. Alternatively, Getty Trust's Categories for the Description of Works of Art (CDWA), which was created to describe the visual arts and related disciplines, also includes rules for best practices and conceptual frameworks for cataloguing, describing, and accessing information about artworks, architecture, and other types of material culture. It also allows for linked data, and uses the Getty vocabularies and descriptive practices.

METADATA DESIGN ACTIVITY

- **Focus on people implementation: Write up a brief description of how this metadata should be managed in the collection. To create metadata records in this schema for this scenario, what training, experience, oversight, or quality control might be needed? What are the best and worst case scenarios of this metadata schema in action? Consider unreliability, safety/social hazards, and gaming the system in addition to good faith attempts to enter in and use the metadata as intended.**

The metadata for this project will be crowd-sourced and entered by users so it will be necessary to provide resources to ensure quality control and standardization. One solution will be to use controlled vocabularies in the value fields to maintain consistency. A dictionary of terms will also be necessary to educate users about the controlled vocabulary and terminology. A training video may also be necessary to educate users how to use the database which will increase familiarization with the database and the metadata entry process. This will be particularly helpful to the user-base, who has a broad range of technology literacy skills and familiarity. Ideally, there would be someone who could provide support and clarity to users as the database becomes populated. The worst case scenario for this database / collection would be searchability issues which correlate to the risk that entries would be incorrectly written, or information would be missing due to the varying level of knowledge that users have when entering information into the database. It will be important to have an educational component to the database, as well as standardization.

METADATA PROCESSING ACTIVITY

- **Make note of what functions didn't work as expected in OpenRefine:**
    - All of the functions for organizing data in OpenRefine were similar to those found in spreadsheet software; however I found the names of the functions were not very intuitive or indicative of what function they perform. For example, the distinctions between clustering, faceting, and filtering.
- What functions covered in the tutorial you'd expect to use most often to:
    - Correct common errors in metadata (such as duplicate records, spelling, blank space,)
        - Transform for removing brackets, colons, semi-colons, etc. in data and trim leading and trailing whitespace
        - Clustering
    - Implement changes in vocabularies (such as replacing a derogatory/discriminatory terms in the "subject" field with a respectful one)
        - Clustering and edit - Merge, selected and re-cluster - (using fingerprint method)
        - The replace function
    - Give quick summaries of items in the collection (such as # of unique creators, most and least common subjects, most and least used items)
        - Facets
        - Sorting