

NYU Moving Image Archiving and Preservation (MIAP) Program
CINE-GT 1803: METADATA FOR MOVING IMAGE COLLECTIONS

Instructor

Shawn Averkamp
sa180@nyu.edu

Course Description

Students in this course will learn about describing and managing moving image collections through metadata, or “data about data”. Metadata may be defined as “structured information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource.” Because it facilitates the access, management and preservation of moving image resources, it is crucial that metadata be created and collected throughout the life cycle of the resource. The creation and use of metadata requires knowledge and experience using various digital tools. To facilitate students’ skills in the practical implementation of metadata within real-world contexts, this class will include investigation of technologies for data storage and exchange, building on the digital literacy class in the first semester. Core concepts will include data modeling, resource description, and databases. Students will become familiar with tools to create and manage metadata.

Learning Objectives

Upon completion of the class students will:

- Understand how metadata supports various functions in the moving image archives
- Understand how to model metadata to support these functions; models allow one to take a broad view of the information needed for access and preservation to moving image resource.
- Understand specific metadata schemes used for describing, providing subject access to and managing moving image resources with an appreciation for the strengths and weaknesses of specific metadata schemes
- Be able to appreciate the importance of standards for describing and preserving moving image resources
- Understand how metadata is implemented and used in a variety of settings.
- Gain experience in creating metadata in different environments and using different tools
- Understand how to evaluate the metadata needs for a particular collection and implement it in a database.

Course Texts

Available online and indicated under each week. The following are texts used throughout the course, along with the articles listed below in the class descriptions and are available on NYU Google Drive.

Metadata. Marcia Lei Zeng and Jian Qin. Chicago: ALA, Neal-Schuman an imprint of American Library Association, 2016. ISBN: 9781555709655.

Other readings can be found in the course Google drive.

Class Topics

Outline of Topics, Assignments

| Week | Topic | Activities and assignments |
|-------------|--|--|
| 1 | Introduction to metadata; metadata models | Activity: Metadata categories, installing text/code editors |
| 2 | Metadata models | Activity: Data modeling exercise |
| 3 | Trip to Culpeper -- NO CLASS | |
| 4 | Data structure standards | Due: Data modeling homework Activity: XML |
| 5 | Descriptive cataloging standards | Due: Final project proposal Activity: XML |
| 6 | Controlled vocabularies, cataloging workshop | Activity: Cataloging workshop, JSON |
| 7 | Linked data | Activity: Wikidata |
| | Spring Break -- NO CLASS | |
| 8 | Application profiles; metadata documentation | Due: Data model Activity: Building an application profile; Wikibase/Wikidata |
| 9 | Metadata interoperability and crosswalks | Activity: creating a metadata crosswalk |
| 10 | Preservation and technical metadata | Due: Final project application profile Activity: Extracting technical metadata |
| 11 | Rights metadata; METS: Migrating and managing metadata | Due: Data mapping assignment Activity: Creating a METS structural map; working session on final project metadata |
| 12 | Metadata quality and remediation, metadata services | Activity: MDQC, OpenRefine |
| 13 | Metadata sharing and reuse | Due: sample records in Wikibase Activity: Querying with APIs, SPARQL |
| 14 | Student presentations | Due: Final assignment |

Assignments

#1: In class assignments and homework. Most classes will include activities that are done either in class or as homework, either as a group or individually. Students will be required to turn in or otherwise demonstrate the results of weekly assignments before the start of the following class meeting, regardless of whether the activity was completed in class or at home as homework. (30%)

(For the following see below under “Formatting” for file naming conventions.)

#2: Data mapping project. Create a crosswalk from our class application profile (YouTube Nighttime routine videos) to two data structure standards (ex., MARC, MODS, **PBCore**, VRACore, Dublin Core, Schema.org, etc.). Map a minimum of 15 fields, selecting fields from different categories of information (descriptive, physical, legal, preservation, technical). Describe strengths and weaknesses of each data standard as it relates to the original class application profile (minimum of 1 paragraph per standard); (20% of final grade)

#3: Final metadata project. Create or repurpose a moving image or other type of collection you can physically or digitally access. You will work toward building and populating a class Wikibase (Wikidata) instance with metadata for this collection throughout the semester. You will start by specifying a data model (what entities you will describe and what are their relationships), and then will build an application profile for the collection, which can borrow properties from established metadata vocabularies, but should be localized for your collection needs. Later you will apply your application profile to Wikibase and create necessary properties to describe your collection items. You will populate the Wikibase with a set of at least 10 records for items in your collection. For the last class, you will give a brief presentation on your collection to the class. The project should demonstrate your understanding of entity relationships, metadata standards and their application, and controlled vocabularies. The project will constitute 40% of the final grade for the class. (40% of final grade)

Grading

Grades will be based on the following:

- Class participation and attendance (10%)
- Class activities/homework (20%);
- Data mapping project (30%)
- Final metadata project (40%)

Feedback on assignments will be given electronically. Unless given an extension by the instructor, late submission may result in a reduction in the grade. Evaluation of the two projects will be based on originality, completeness, accuracy (grammar, spelling), and timeliness.

Week 1 -- Introduction to metadata

Due this class:

- Reading: Gilliland, Anne. (2016). "Setting the Stage," in Introduction to Metadata, 3rd ed. <https://www.getty.edu/publications/intrometadata/setting-the-stage/>
- Reading: "Introduction" in Metadata. Pages 3-22 Available on NYU Google Drive https://drive.google.com/file/d/1DbSXy6h6mLoEVLeeZ5ptz9eA_8n-lkW/view?usp=sharing
- Reading: Doctorow, Cory. (2001). "Metacrap: Putting a torch to the seven straw-men of the meta-topia," <http://www.well.com/~doctorow/metacrap.htm>
- Reading: Onuoha, Mimi (2016). "The Point of Collection," Data & Society Points, <https://points.datasociety.net/the-point-of-collection-8ee44ad7c2fa>

Topics/activities:

- Overview of class goals and expectations; review of syllabus.
- Overview of the principles of metadata
- Introduction to metadata models
- Activity: Metadata categories
- Activity: Installing code/text editors and intro to data formats (CSV, XML, JSON)

Week 2 -- Metadata models

Due this class:

- Reading: Steve Höberman, “Section 1: Data Modeling Introduction” and “Section 2: Data Model Components”, Data Modeling Made Simple, 2nd edition. Technics Publications, 2012. Online version available from NYU Libraries: <https://getit.library.nyu.edu/go/9442309>
- Reading: Tillett, Barbara. “What is FRBR?” Washington, D.C.: Library of Congress, Cataloging Distribution Service, 2004. <https://www.loc.gov/cds/downloads/FRBR.PDF>
- Reading: Carlyle, Allyson. “Understanding FRBR As a Conceptual Model: FRBR and the Bibliographic Universe.” Library Resources & Technical Services, 2006, 50(4), 264-273. <http://dx.doi.org/10.5860/lrts.50n4.264> (read up to “Other Cataloging Models”).
- Reading: MovieLabs. “White Paper -- Creative Works Ontology for the Film and Television Industry.” September 2018. <https://movielabs.com/creative-works-ontology/> (Read the white paper at <https://movielabs.com/wp-content/uploads/2018/09/A-Creative-Works-Ontology-for-the-Film-and-Television-Industry-Final-2018-9-24.pdf>)
- Reading: Van Malssen, Kara. “Bibframe AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources.” (submitted May 15, 2014). <http://www.loc.gov/bibframe/docs/pdf/bibframe-avmodelingstudy-may15-2014.pdf> (Read p. 2-43)

Topics/activities:

- Metadata models for library, archive and museum contexts
 - Why do we need metadata models?
 - How are data models and ontologies built?
 - Understanding user needs and how they affect data models
 - FRBR (Functional Requirements for Bibliographic Records) and FRBR’s applicability to moving image materials. <http://www.ifla.org/VII/s13/frbr/frbr.htm>
 - BIBFRAME model: <https://www.loc.gov/bibframe/docs/bibframe2-model.html>
 - PBCore model: <https://pbcore.org/data-model>, <https://pbcore.org/elements>, <https://pbcore.org/attributes>
 - Cinematographic Work Standard (EN 15907): http://filmstandards.org/fsc/index.php/EN_15907
 - MovieLabs ontology
 - Introduction to linked data
- **Activity:** Understanding data models: FRBR, BIBFRAME, EN 15907
- **Activity:** User needs and data modeling exercise
- **Homework:** Diagramming MovieLabs ontology

Week 3 -- Trip to Culpeper -- NO CLASS

Week 4 -- Data structure standards and element sets

Due this class:

- Homework on conceptual data modeling
- Review: Riley, Jenn. "Seeing Standards: a Visualization of the Metadata Universe"
<http://jennriley.com/metadatamap/>
- Reading: Metadata Pages 402-420; 445-456.
- Library of Congress (2009). "Understanding MARC Records: What is a MARC Record and Why is it Important?" Library of Congress, 2009. <http://www.loc.gov/marc/umb/um01to06.html>

Topics/activities:

- What is a data structure? Schemas and rules
 - Structure vs content vs value standards
- Data Structures for libraries, archives, museums:
 - Discuss and compare data structures: [MARC21](#), [MODS](#), [Dublin Core](#), [PBCore](#), [EBU Core](#), [EAD](#)
- Review metadata records
- Activity: XML 9 Bring your laptops and have a code editor such as [Sublime Text](#) or [Visual Studio](#) installed.)
 - XML with Lynda.com:
<https://www.lynda.com/XML-tutorials/XML-Essential-Training/661763-2.html> (Parts 1, 2, 3.1, 4.1, 5)
 - PBCore Cataloging Tool (creating PBCore XML): <https://pbcore.org/cataloging-tool>
 - PBCore XML validator: <https://pbcore.org/validator>

Week 5 -- Descriptive cataloging standards (content standards)

Due this class:

- Final project proposal
- Reading: Describing Archives: a Content Standard (DACS). Chicago: Society of American Archivists, 2013, 2nd edition,
http://www2.archivists.org/standards/DACS/statement_of_principles
- Reading: FIAF Moving Image Cataloging Manual. International Federation of Film Archives, 2016, especially p. 1-11. Available at:
<https://www.fiafnet.org/images/tinyUpload/E-Resources/Commission-And-PIP-Resources/CDC-resources/20160920%20Fiaf%20Manual-WEB.pdf>
- Miller, L. (2011). Resource Description and Access (RDA): An Introduction for Reference Librarians. Reference & User Services Quarterly, 50(3), 216-222.
<http://www.jstor.org.proxy.library.nyu.edu/stable/41241166>

Review:

- RDA Toolkit: Introduction. <http://access.rdatoolkit.org/> (sign up for a 30-day free trial to access). Section 0. Also look at RDA to MARC mapping under Tools tab.

Topics/activities:

- Review descriptive standards: AACR2 (Anglo-American Cataloging Rules, 2nd Edition), Resource Description and Access (RDA), FIAF Cataloging Manual, Describing Archives: a content standard (DACS)
- Activity: Cataloging a video with FIAF
--https://archive.org/details/0749_Its_a_Cats_Life_08_12_50_00
- Activity: more XML (bring your laptops)

Week 6 -- Cataloging Workshop; controlled vocabularies and authority control

Due this class:

- Reading: Harpring, Patricia. Introduction to Controlled Vocabularies, Terminology for Art, Architecture, and Other Cultural Works. Chapter 2: “What Are Controlled Vocabularies?” and Chapter 3: “Relationships in Controlled Vocabularies.” Los Angeles, CA: J. Paul Getty Trust, 2010. http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/wat.html and http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/relationships.html
- Library of Congress (2004). “Understanding MARC Authority Records: What is a MARC Record and Why is it Important?” Library of Congress, 2004. <https://www.loc.gov/marc/uma/pt1-7.html>

Review:

- Library of Congress Linked Data Service. <http://id.loc.gov/>
- Library of Congress Moving Image Genre-Form Terms http://www.olacinc.org/sites/default/files/Genre-Form-Headings_2018_1.pdf and Library of Congress Moving Image Genre/Form Headings H 1913: <http://www.loc.gov/catdir/cpsd/h1913.pdf>
- EAC-CPF / SNAC: <http://snaccooperative.org/>
- VIAF: <http://viaf.org>
- Getty vocabularies: <http://www.getty.edu/research/tools/vocabularies/>

Topics/activities:

- Controlled vocabularies and authority control
- Activity: Cataloging exercise (bring your laptops)
- Activity: JSON

Week 7: Mar. 14 -- Linked data; identifiers

Due this class:

- Reading: Berners-Lee, Tim, Hendler, James, and Lassila, Ora, (2001). “The semantic web.” Scientific American, May 2001, p. 29-37. <https://www-scientificamerican-com.proxy.library.nyu.edu/article/the-semantic-web/>
- Reading: Tillman, Ruth (2016). “An Introduction to RDF for Librarians (of a Metadata Bent).” <http://ruthillman.com/introduction-rdf-librarians-metadata/>
- Reading: EIDR. “About EIDR.” (including FAQs) <https://eidr.org/about-us/>
- Reading: Fauconnier, Sandra, (2018). “Many faces of Wikibase: Rhizome’s archive of born-digital art and digital preservation.” <https://wikimediafoundation.org/2018/09/06/rhizome-wikibase/>
- Reading: Wikibase/DataModel/Primer: <https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer>
- Review: Dublin Core Metadata Initiative (DCMI) terms: <http://dublincore.org/documents/dcmi-terms/>
- Review: VIAF: <http://viaf.org>
- Take the Wikidata Tours: <https://www.wikidata.org/wiki/Wikidata:Tours>

Topics/activities:

- RDF and linked data
- Introduction to Wikidata
- Activity: Wikidata editing

Week 8 -- Application profiles; metadata documentation

Due this class:

- Homework: Final project data model

- Reading: Heery, Rachel and Patel, Manjula (2000). "Application profiles: mixing and matching metadata schemas." *Ariadne* 25.
<http://www.ariadne.ac.uk.proxy.library.nyu.edu/issue/25/app-profiles/>
- Review: DLF AIG Metadata Application Profile Clearinghouse:
<https://dlfmetadataassessment.github.io/MetadataSpecsClearinghouse/>
- Review: Metadata application profiles and data dictionaries:
 - Carnegie Hall: <https://github.com/CarnegieHall/digitalcolls-metadataprofile>
 - University of Nebraska-Lincoln: <https://unl.libguides.com/c.php?g=813899&p=5807669>
(Review each tab)
 - New York Art Resources Consortium (NYARC):
<https://www.nyarc.org/sites/default/files/web-archiving-profile-version2.pdf>
 - NISO data dictionary--technical metadata for digital still images:
https://groups.niso.org/apps/group_public/download.php/14697/z39_87_2006_r2011.pdf (in particular, sections 5 and 6)

Topics/activities:

- PBCore -- <http://pbcore.org/>
- Application profiles/data dictionaries/Wikidata properties
- Activity: [Building an application profile](#)

Week 9 -- Metadata interoperability and crosswalks

Due this class:

- Reading: Metadata: p. 347-357. Available from NYU Google Drive:
<https://drive.google.com/open?id=1F66graOhSgyJZ9Iz8EyPcjWkRbPk7z1k>
- Reading: Woodley, Mary S., revised by Baca, Murtha (2016). "Metadata Matters: Connecting People and Information" in Introduction to Metadata, 3rd ed.
<http://www.getty.edu/publications/intrometadata/metadata-matters/>
- Reading: Marcus, Cecily and Carlson, Sarah (2018). "Out of the Shadows: Bringing African American Digital Collections Together in Umbra Search African American History." Open Library of Humanities, 4(2), 17. <http://doi.org/10.16995/olh.279>
- Reading: [PBCore Mappings](#)

Topics/activities:

- How do different metadata standards work together? Explore issues about metadata interoperability
- Creating and understanding crosswalks
- Wikidata properties
- Activity: Creating a metadata crosswalk
- Activity: Properties in Wikibase/Wikidata
- Homework: Data mapping assignment

Week 10 -- Preservation and technical metadata

Due this class:

- Homework: Final project application profile
- Reading: Caplan, Priscilla and PREMIS Editorial Committee. Understanding PREMIS, revised 2017. <https://www.loc.gov/standards/premis/understanding-premis-rev2017.pdf>
- Reading: Tactical Technology Collective, "Behind the data: investigating metadata." *Investigating the Invisible*. <https://exposingtheinvisible.org/guides/behind-the-data-metadata-investigations/>
- Reading: Ng, Yvonne, "What is metadata? Our new video explains."
<https://blog.witness.org/2014/10/video-metadata/>

- For the following, get a general familiarity with them:
 - Review: PREMIS 3.0 <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
 - Review: SMPTE standards <https://www.smppte.org/top-standards> and <https://www.smppte.org/standards/document-index>

Topics/activities:

- Review data requirements and standards for preservation and technical metadata, including sources for controlled vocabularies for terms
- Generating metadata from digital objects and data management systems
- Embedded metadata
- Activity: Tools for working with technical metadata
- Activity: [PREMIS](#)

Week 11 -- Rights metadata; Structural and package metadata

Due this class:

- Homework: Assignment on data mapping
- Reading: Whalen, Maureen. "Rights Metadata Made Simple." In Introduction to Metadata. http://www.getty.edu/research/publications/electronic_publications/intrometadata/rights.html
- Review: Traditional Knowledge (TK) labels. <http://localcontexts.org/tk-labels/>
- Reading: Kirby, J. (2017) "An Introduction to Traditional Knowledge Labels and Licenses." Commons Knowledge Blog. <https://publish.illinois.edu/commonsknowledge/2017/09/07/an-introduction-to-traditional-knowledge-labels-and-licenses/>
- Review: Rightsstatements.org <https://rightsstatements.org/en/>
- Reading: METS Overview <http://www.loc.gov/standards/mets/METSOverview.v2.html>
- Hardesty, Julie. (2017). "The Semantics of Metadata: Avalon Media System and the Move to RDF." Code4Lib Journal Issue 37 <https://journal.code4lib.org/articles/12668>

Topics/activities:

- Rights and permissions metadata
- Container formats and strategies for packaging/connecting metadata and content
- Review of RDF, PREMIS, and identifiers
- Metadata Encoding and Transmission Standard (METS)
 - What is a METS document?
 - How/where is METS used?
 - Alternatives to METS XML
- Activity: Working session on final project metadata

Week 12 -- Metadata quality and remediation, metadata services

Due this week:

- Reading: Bruce, Thomas and Hillman, Diane (2004) "The Continuum of Metadata Quality: Defining, Expressing, Exploiting." Published in "Metadata in Practice," ALA Editions. <https://ecommons.cornell.edu/handle/1813/7895>
- Review: DLF AIG Metadata Working Group Metadata Assessment Toolkit: <https://dlfmetadataassessment.github.io/>
- Reading: Metadata. Pages 251-258 <https://drive.google.com/open?id=13q3Clx3RMNHkaeTYGvw8UXkyjK9bQxky>

Topics/activities:

- Metadata quality and dealing with messy data

- Activity: MDQC
- Activity: cleaning up metadata (Google sheets)

Week 13 -- Metadata sharing and reuse

Due this class:

- Homework: Final project records in Wikibase
- Open Knowledge Foundation. “What is Open Metadata?” Published in “Open Metadata Handbook,” Wikibooks (accessed January 13, 2019).
https://en.wikibooks.org/wiki/Open_Metadata_Handbook/Open_Metadata
- Always Already Computational (2018). “The Santa Barbara Statement on Collections as Data.”
<https://collectionsasdata.github.io/statement/>
- DPLA, “API Codex” <https://pro.dp.la/developers/api-codex> (Read at minimum: API Codex, API Basics, Requests (skim this), Responses, Philosophy)

Topics/activities:

- Metadata licenses
- Methods for sharing metadata records and digital assets
 - OAI-PMH
 - APIs
 - Data dumps (ex. <http://americanarchive.org/help/obtain-metadata>)
 - IIIF
 - Frictionless data and JSON
- Activity: Using APIs and SPARQL to query and harvest metadata

Week 14 -- Student presentations

Due this class:

- Final project

Topics/activities:

- Student presentations

CINE-GT 1803: Final Project

Create or repurpose a moving image or other type of collection you can physically or digitally access. You will work toward building and populating a class Wikibase (Wikidata) instance with metadata for this collection throughout the semester. You will start by specifying a data model (what entities you will describe and what are their relationships), and then will build an application profile for the collection, which can borrow properties from established metadata vocabularies, but should be localized for your collection needs. Later you will apply your application profile to Wikibase and create necessary properties to describe your collection items. You will populate the Wikibase with a set of at least 10 records for items in your collection. For the last class, you will give a brief presentation on your collection to the class. The project should demonstrate your understanding of data modeling for a specific domain and user base, entity relationships, metadata standards and their application, and controlled vocabularies. The project will constitute 40% of the final grade for the class.

- Choose a collection to use for the project.
- Create a conceptual data model for the collection -- what entities you are describing and their relationships.
- Create an application profile to document the elements you need and their specifications (e.g. repeatable, mandatory) including use of any controlled vocabularies.
- Create entity types and properties in the class Wikibase instance based on your application profile.
- Create at least 10 records for resources from your collection. Also create records as appropriate for any linked entities.

For each part of this project, send your submission by email to sa180r@nyu.edu. Name your files as follows:

18s_1803_Lastname_A2_datamodel

18s_1803_Lastname_A2_appprofile

18s_1803_Lastname_A2 (for final assignment)

Note that submission of the CA profile will be a URL to your system.

1. Create or select a collection. Create or select a collection – any collection of objects or information packages that you can physically or digitally access. The collection of objects should have some sort of enduring value either to you personally or to the human record. Identify your collection and describe why it has enduring value. (You don't need to identify every item in your collection, just include a description of the types of items that will be found in it.) Describe the potential users of your collection and explain how they will want to navigate it and access its items. Students often choose their own personal collections (e.g. their own home movies, music libraries, objects of importance) or collections related to their internships (assuming the project

doesn't duplicate work done on the internship). This proposal can be brief: a couple of paragraphs to 1 page.

Proposal due: Feb. 28, 2019

2. Create a data model for the collection. After assessing your users and their needs, write a set of business rules for your data model. This may be in narrative or list form and should describe the entities or things in your domain, their relationships to each other, and the requirements and restrictions of each. Using data modeling techniques create a graphic data model that shows the entities that you will describe and the relationships between them. You may use a diagramming tool (such as [Lucidchart](#) or [Gliffy](#)) or you could draw it by hand and take a picture of it. Use entity-relationship diagram conventions to show relationships (including cardinalities, i.e. one-to-many, etc.).

For the data model you do NOT have to include every data element that you will use for each entity, since that is part of creating the application profile.

3. Create an application profile to indicate the metadata properties (aka elements) you need and their specifications. In your data model you diagrammed the entities and relationships in your collection domain, but in the application profile you will detail the properties and relationships of those entities in a document. The first step is to determine the pieces of information that you deem are important information about the items in your collection, which could include descriptive, technical, or preservation-related metadata. This will depend on the user needs for your collection is: e.g. for searching the items in the collection, for preserving them, for managing their location. Specify whether each is required/optional repeatable/not repeatable and the data type of each, i.e. string, integer, date, controlled vocabulary, URI. If the element requires a controlled vocabulary specify which vocabulary you will use (if an established one) or provide your local controlled values. Organize the profile according to the entities in your data model with separate sections for each entity and properties that apply to each. Make sure to include a key to help readers interpret your profile. (You can use the [sample application profile](#) as your guide.) If you are borrowing or refining properties from other vocabularies, including Wikidata, note the equivalency in a separate field in each property's table.

Give an introduction (1-2 paragraphs) at the beginning of your application profile refining your collection proposal, explaining the scope and content of your collection and why you chose and describing your target users and their specific search and access needs. Give at least two examples of how properties you've chosen for your profile will help meet those needs.

4. Create at least 10 sample records in the class Wikibase describing items in your collection. You will first need to create entity types and properties from your application profile in Wikibase before you can create Wikibase records ("items"). We will learn how to query your records and your classmates' with SPARQL in class on May 2, so please ensure you have your records entered before class.

5. Present your project to the class. Each student will give a presentation about the project, between 5-10 minutes. You will introduce your collection, why you chose it, the target audience and their needs, how you modeled it, and a brief explanation of the elements and controlled vocabularies you used to describe the collection and related entities. Plan to share a few sample records and share any thoughts on how you would ensure one or two dimensions of quality as well as your experience completing this assignment.

Presentation due with final assignment.

Your final project will consist of the following components:

1. Collection description and value (revised as appropriate from your proposal). You can include this in the introduction to your application profile.
2. Data model business rules and diagram
3. Application profile using the template provided or one of your own design. (Make sure you remember to include your Wikibase properties and ids as the labels of each entry.)

Note that the drafts are due throughout the semester so that I can give you feedback. Please stick to the schedule so that you are able to complete the assignment on time!

Grading of assignments

The following factors will be taken into account:

- **Content/accuracy.** You've chosen appropriate elements, defined controlled vocabularies where needed, created accurate metadata records that conform to the profile, written the profile in a clear and accurate way (with a minimum of typos and/or grammatical errors). 40% of assignment grade.
- **Completeness.** You've included all elements of the assignment described above. 30% of assignment grade.
- **Originality/complexity.** You show originality in your choice of collection, explanation of user requirements, and use of elements/controlled vocabularies. 10% of assignment grade
- **Presentation.** You give a presentation that covers the components of the project in a clear and understandable way. 10% of assignment grade
- **Timeliness.** Your assignment (including drafts due at different stages) is on time. 10% of grade.

Week 2 -- Class activities

Activity 1: Data model review

Break into 3 small groups. Each group will review an established data model: [FRBR](#), [BIBFRAME](#), or Cinematographic Work Standard ([EN 15907](#)). Work together to answer the following questions. (You may need to look at other parts of the data model's website or search the internet to answer some of these.):

- What user needs does this model try to serve?
- What is the domain of the model? What is the scope? (What does the model claim to cover? What does it not appear to cover?)
- What are the entities in the model? Give a brief definition of each.
- Who created or maintains the data model?

Choose as many group members as you like to present your review to the class.

Activity 2: Data modeling exercise

You have been asked to overhaul the catalog/digital asset management system and access platform for a video archive. You have been given free reign to interview whoever you need to build the most useable archive possible. Who will you interview? How will you learn about the archive's users' needs? How will you design a data model that supports as many of these needs as possible?

1. Select an archive and from the list below. Describe the content that the archive provides access to. Is it physical or digital? Both? What are the characteristics of the content? What else can you learn about how the content is organized? Do items belong to collections? How are items and collections described? How does the video archive relate to the larger archive? How does the archive provide access? What are other services the archive offers? (15 minutes)
2. Based on what you can learn from the website about the archive's mission and content, brainstorm all possible users of the system and list them on post-its. For each, list some anticipated needs for the future system. Organize your user post-its on a table or the wall in order of priority. (10 minutes)
3. Taking turns, each person in your small group will play the role of one of your identified user personas. Interview this person to better understand how they want to use your archive, how they'll be searching, what limitations they have. As you're listening, take note of the entities and properties your system and data model will need to support to meet your users' needs. (10 minutes)
4. Review your notes and list out all of the needs/requirements on post-its and list below each the user types to which they apply. Considering your user type priority, now prioritize these requirements. (10 minutes)
5. Based on everything you've learned about the archive and its users, draft some brief business rules for your archive and draw a conceptual data model listing all of the entities that will need metadata (properties) and the relationships between them. (You don't need to list the properties in this diagram, but you can include them in preliminary sketches to help understand

what properties should be “strings” and which should be things (other entities). Be sure to consider your archive’s relationship with the larger archive, if applicable, and include those entities and relationships as appropriate. (20 minutes)

Video Archives:

- Film Noir Foundation <http://www.filmnoirfoundation.org/home.html>
- George Balanchine Video Archives <http://balanchine.org/balanchine/03/gbfvideoarchives.html>
- Science Friday <https://www.sciencefriday.com/videos/>
- Chicago PD Crime Archives
<https://home.chicagopolice.org/office-of-communications/crime-video-archives/>
- UNC Dept. of Medicine Grand Rounds Video Archive
<https://www.med.unc.edu/medicine/events/grand-rounds/video-archives>
- Live with Kelly and Ryan Video Archive <https://kellyandryan.com/videos/>
- City of Carson, California Video Archive <http://ci.carson.ca.us/Government/videoarchive.aspx>

Week 2 -- Assignment: Diagramming MovieLabs Ontology

Review our reading of the MovieLabs Ontology whitepaper (<https://movielabs.com/creative-works-ontology/>) and try to diagram the ontology’s entities and relationships. (Diagram the entities and relationships only--don’t worry about the attributes. Only “things,” no “strings.”) Include any subclasses of entities using the relationship “subClassOf”. You may use whatever methods for drawing the diagram you like, using online tools such as Lucidchart or Gliffy or drawing on paper and taking a picture. Following the exercises we completed in class, use ovals for entities and arrows for relationships.

Week 10 -- Class activity

Break out into three small groups to assess and record preservation metadata for one of the following digital objects:

Group 1: Audio -- Passamaquoddy vocabulary items:

<https://www.loc.gov/item/2015655558/>

Zip file: https://drive.google.com/open?id=1ns7fX3Yf6TxCad_4uVYvtAn6RiwdVRYC

Group 2: Image -- Mrs. Charles Benning and baby in their shack home at Shantytown, Spencer, Iowa.

<https://digitalcollections.nypl.org/items/70ceaeb0-b383-0136-911f-75bd8d41312b>

Zip file: https://drive.google.com/open?id=1GJN56-gxsk7AdlOTF_J5NoGdu9yct9Bi

Group 3: Video -- President Harding and Calvin Coolidge:

<https://www.loc.gov/item/mp76000342/>

Zip file: https://drive.google.com/open?id=1xNwpO0BQQANk0kY2_pxNi3wrVoh_jw7E

Use MediaInfo or ExifTool, whichever is appropriate for your file types, to view the technical metadata. Using the technical metadata and available metadata from the digital object's webpage, try to fill out as much as you can about the Objects (files, representations, intellectual entity) and create metadata for one Event in a copy of the [PREMIS metadata template](#). Use the [PREMIS Data Dictionary](#) as a reference for filling in the elements.