

MOVING IMAGE ARCHIVING & PRESERVATION PROGRAM

CINE-GT 1803: METADATA FOR MOVING IMAGE COLLECTIONS

This will be available on NYU Classes and updated periodically.
Spring 2016 – Mondays, 6 pm – 10 pm, 721 Broadway room 652
Instructor: Rebecca Guenther

Email: rg153@nyu.edu
cell.: (703) 298-0157

Office hours: Mondays 4-6pm by appointment. 721 Broadway room 679
Class dates: Jan. 25; Feb. 1; Feb. 8; Feb. 29; Mar. 7; Mar. 21; Mar. 28; Apr. 4; Apr. 11;
Apr. 18; Apr. 25; May 2; May 9 (makeup class for Feb. 22 to be determined; may be field trip to MoMA)

GOALS:

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. Topics include how metadata supports various functions in the moving image archives; metadata models; specific metadata schemes used for describing, providing subject access to and managing moving image resources; the importance of standards for resource description; information needed for preservation of moving image resources; and how metadata is implemented and used in a variety of settings.

In order to facilitate students’ skills in the practical implementation of metadata within real-world contexts, this class will have an added component of technologies for data storage and exchange. Core concepts will include data modeling, data quality, and databases. Students will also become familiar with tools to create and manage metadata.

EXPECTATIONS:

Attendance at all classes is expected unless excused, as our work together will be intensive. Grades will be based on a combination of class preparedness and participation (20%); metadata searching comparison (10%); data mapping project (30%); final metadata project (40%). Feedback on assignments will be given electronically.

TEXTS: The following are texts for the course, along with the articles listed below in the class descriptions. Books are on reserve at Bobst. Also available on NYU Google Drive.

- *Metadata*. Marcia Lei Zeng and Jian Qin. New York, NY : Neal-Schuman Publishers, Inc., 2008 ISBN: 9781555706357
Available from NYU Google Drive at:

- <https://drive.google.com/a/nyu.edu/folderview?id=0B15icbsejHfMdFdLa2w3Wm-pwTms&usp=sharing>
- *Descriptive Metadata for Television: an End-to-End Introduction*. Mike Cox, Linda Tadic, Ellen Mulder. Amsterdam: Focal Press/Elsevier, 2006. ISBN: 0240807308
Available from NYU Google Drive at:
<https://drive.google.com/a/nyu.edu/folderview?id=0B15icbsejHfMcGJhc3ZteDN-VXzg&usp=sharing>

TECHNOLOGIES

- FileMaker Pro
- oxygen XML editor

ASSIGNMENTS

#1: Metadata for searching moving image collections comparison. Choose 2 of the websites from the list and try searching for AV material of interest to you. Search for at least two types of resources at each (e.g. involving a person, subject, type of material, etc.) that you are interested in. Answer the questions detailed in the assignment description. (See longer description on NYU Classes). Due March 7 (10%).

#2: Data mapping project. Create a crosswalk between three data structure standards. Map a minimum of 20 fields, selecting fields from different categories of information (descriptive, physical, legal, preservation, technical). Describe strengths and weaknesses of each data standard (minimum of 1 paragraph per standard); See longer description on NYU Classes). Due April 11 (30%)

#3: Metadata project. Analyze a moving image collection you can physically or digitally access. Throughout the semester, you will work toward building and populating a database for this collection. You will start early in the semester with creating a data model and application profile for the collection, which will be based on an existing metadata standard, but localized for your collection needs. Later you will build your database in FileMaker Pro, including a search/data entry user interface. Finally, you will populate the database with a set of sample records. The database should demonstrate your understanding of entity relationships, data types, controlled vocabularies, and the relationship between local data stores to metadata standards. Each student will give a presentation about his/her project (see longer description on NYU Classes).
Proposal for collection due Feb. 8. Physical model of database due Mar. 28. Database setup due Apr. 4. Database design first draft due Apr. 18. Database design final draft Apr. 18. Database layout final draft May 2. Final assignment due May 9 (40%)

MIAP DIGITAL ARCHIVE: Your papers will be made part of the MIAP Digital Archive in a private space for faculty use, and on the MIAP web site, where appropriate. Please inform me of any papers that cannot be published on the web due to confidentiality restrictions or special circumstances. In some cases, the title of a paper will be published, but access to the paper will be restricted to selected MIAP faculty and staff. File submission format for assignments:

year semester_class number_author's last name_assignment number.ext

FIELD TRIP

MoMA Film Study Center: Date TBD

Outline of topics, activities and assignments (makeup class TBD)

	Date	Topic	Activities and assignments
1	1/25/16	Introduction to metadata	Exercise: Introduction to metadata
2	2/1/16	Metadata models	Exercise: Data models Exercise: ER notation
3	2/8/16	Data structure standards; Intro to databases	Workshop on databases (Deena Engel) Due: Proposal for collection
4	2/29/16	Data structure standards; descriptive and archival standards	Exercise: comparison of EAD, MARC 21 and MODS Due: Homework on databases (ungraded)
5	3/7/16	Metadata syntaxes; XML	Exercise: Creating XML Due: Assignment #1: Using metadata for searching
6	3/21/16	Controlled vocabularies and authority control; Databases	Workshop on databases (Deena Engel)
7	3/28/16	Metadata systems and tools; Databases	Workshop on databases (Deena Engel) Due: Physical model of final project database
8	4/4/16	Metadata interoperability; XSLT	XSLT Workshop (Ashley Blewer) Due: Initial FMP database setup
9	4/11/16	Cataloging workshop	Workshop by Andrea Leigh Due: Assignment #2: Data mapping
10	4/14/16	Migrating and managing metadata	MAKE UP CLASS Guest: Thelma Ross
11	4/18/16	Preservation, technical and legal data	Exercise: Technical metadata Due: Initial database setup
12	4/25/16	Managing and using metadata	Workshop on Collective Access (Seth Kaufman) Due: FMP database design draft
13	5/2/16	Linked Data	Workshop on Linked Data tools Due: FMP database layout draft
14	5/9/16	Student presentations	Due: Final assignment

Class 1: January 25 – Introduction to Cataloging and Metadata

Due this class:

- Gilliland, Anne. (2008). "Setting the Stage," in *Introduction to Metadata*. http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html
- Gill, Tony (2008). "Metadata and the Web," in *Introduction to Metadata*. http://www.getty.edu/research/publications/electronic_publications/intrometadata/metadata.html
- Reading: *Descriptive Metadata for Television*. Pages 1-18 (on reserve) Also available on NYU Google Drive.

Topics/activities:

- Overview of class goals and expectations; review of syllabus.
- Overview of the principles of cataloging and metadata
- Review of typical issues with description of different formats and genres
- Compare item-level and collection-level records, and finding aids
- Activity: what is metadata?
- Download and set up software for the class (Oxygen, Filemaker Pro)
- Introduction to final project – students begin thinking of collections they can work with to create and populate a database using FileMaker Pro

Class 2: February 1 – Metadata models

Due this class:

- Reading: Steve Höberman, "Section I: Data Modeling Introduction", *Data Modeling Made Simple*, 2nd edition. Technics Publications, 2012. **Online version available from NYU Libraries:** <https://getit.library.nyu.edu/go/9394338>
- **OPTIONAL:** Reading: Scott Wambler, "Data Modeling 101" Agile Data, <http://www.agiledata.org/essays/dataModeling101.html>
- Reading: Riva, Pat. *Introducing the Functional Requirements for Bibliographic Records and Related IFLA developments* (ASIS&T Bulletin, August/September 2007) <http://www.asis.org/Bulletin/Aug-07/riva.html>
- Reading: *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*, Washington, D.C.: Library of Congress, 1012, p. 3-15 <http://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>
- Review: Metadata Standards for Cinematographic Works (filmstandards.org) http://filmstandards.org/fsc/index.php/Main_Page especially data model: http://filmstandards.org/fsc/index.php/EN_15907 and EN 15744: "Film identification — Minimum set of metadata for cinematographic works", http://filmstandards.org/fsc/index.php/EN_15744

Topics/activities:

- Metadata models for library, archive and museum contexts
 - Define FRBR (Functional Requirements for Bibliographic Records) and FRBR's applicability to moving image materials. <http://www.ifla.org/VII/s13/frbr/frbr.htm>
 - DCMI Abstract Model
 - EN 15907 (*Film identification - Enhancing interoperability of metadata - Element sets and structures*)
 - Introduction to BIBFRAME

- Data models: Entities and Relations
 - Notation (e.g. crows foot, UML)
- Activity: Conceptual Data modeling exercise
 - Hands-on group activity modeling entities and relations
- Activity: Simple entity-relationship diagram

Class 3: Feb. 8 – Data structure standards and element sets (part 1: general);

Introduction to databases

Due this class:

- Proposed collection for final project (submit by email)
- Reading: *Understanding MARC Bibliographic*: <http://www.loc.gov/marc/umb>
- Reading: Data Standards List [on NYU Classes]
- Reading: *Metadata*: p. 15-42; p. 76-81 (Available from Drive, see above)
- Reading: *Descriptive Metadata for Television*. Pages 19-22; 37-54; 113-130 (sample records) (Available from Drive, see above)
- Reading: Steve Hoberman, “Section II: Data Model Components”, *Data Modeling Made Simple*, Technics Publications, 2009. Online version available from NYU Libraries <https://getit.library.nyu.edu/go/9394338>
- Review: Riley, Jenn. “Seeing Standards: a Visualization of the Metadata Universe” <http://www.dlib.indiana.edu/~jenlrile/metadatamap/>
- Review: Metadata Standards for Cinematographic Works (filmstandards.org) EN 15744: “Film identification — Minimum set of metadata for cinematographic works”, http://filmstandards.org/fsc/index.php/EN_15744
- **Handouts:** See handouts on individual data standards **from NYU Classes** (in Handouts folder under Resources/Readings)

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, VRACore
- Review metadata records
- Introduction to data storage using databases (Guest speaker: Deena Engel)
 - Comparison between databases and spreadsheets
 - Database structures: tables/rows/columns, relationships, data types, keys (primary/foreign)
 - Data types
 - Character lengths
 - Query languages, e.g. SQL
 - Database vs UI
 - Examples: MySQL (web application backend), FileMaker (local storage, combining database and UI into one tool)

*****NO CLASS Feb. 15 President’s day *****

*****NO CLASS Feb. 22 (in Culpeper)*****

Class 4: Feb. 29 -- Data structure standards and element sets (part 2: moving image standards); Descriptive and archival content standards

Due this class:

- Homework assignment on databases (ungraded)
- Reading: *Describing Archives: a Content Standard (DACS)*. Chicago: Society of American Archivists, 2007. p. xi-xv [**on NYU Classes under Resources/Readings**].
- Reading: *Metadata*, p. 52-59. (Available from Drive; see above)
- Reading: Digital video archives: managing through metadata (2002) <http://www.clir.org/pubs/reports/pub106/video.html>
- Reading: Steve Hoberman, "Section III: Subject Area, Logical, and Physical Data Models", *Data Modeling Made Simple*, Technics Publications, 2009. **Online version available from NYU Libraries** <https://getit.library.nyu.edu/go/9394338>

Topics/activities:

- Review moving image standards: PBCore, EBU Core, DMS-1, FIAT
- Review descriptive standards: AACR2 (Anglo-American Cataloging Rules, 2nd Edition), Resource Description and Access (RDA), AMIM2 (Archival Moving Image Materials: A Cataloging Manual), Explore RDA models: content, carrier, media; application of FRBR
- Review archival standards: DACS (Describing Archives: a Content Standard) and EAD (Encoded Archival Description)
- Activity: comparison of EAD and MARCXML and MODS

Class 5: Mar. 7 – Metadata syntaxes; XML

Due this class:

- Reading: Myer, Tom. "A Really Really Really Good Introduction to XML", August 2005, <http://www.sitepoint.com/really-good-introduction-xml/>
- Reading: *Metadata*: p. 131-145. (Available from Drive, see above)
- Reading: *Descriptive Metadata for Television*. Pages 76-88. (Available from Drive, see above)
- **ASSIGNMENT #1: Metadata for searching moving image collections comparison**

Topics/activities:

- XML basics
- Schemas: structures and semantics
- Metadata creation and conversion tools
- Identifiers and identification
- Introduction to data storage using XML
 - Role of XML: Data exchange (why most standards have an XML schema) or display (e.g. EAD web access)
- Exercise: Creating XML metadata records using Oxygen

***** NO CLASS MONDAY MARCH 14 (Spring Recess) *****

Class 6: Mar. 21-- Controlled vocabularies and subject analysis; Authority Control ; Databases Part 2 (Deena Engel)

Due this class:

- Reading: Andreano, Kevin. "The Missing Link: Content Indexing, User-Created Metadata, and Improving Scholarly Access to Moving Image Archives." *The Moving Image* 7:2 (Fall 2007), p. 82-99. **[on NYU Classes]**
- Reading: Harpring, Patricia. *Introduction to Controlled Vocabularies, Terminology for Art, Architecture, and Other Cultural Works*. Chapter 2: "What Are Controlled Vocabularies?" Los Angeles, CA: J. Paul Getty Trust, 2010. http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/what.html
- Reading: Churcher, Clare, *Beginning Database Design: From Novice to Professional*, Springer, 2012, Chapters 1 and 2. **Online version available through NYU Libraries:** <http://ezproxy.library.nyu.edu:3926/toc.aspx?bookid=54142>
- Review:
 - LCSH (Library of Congress Subject Headings) (<http://authorities.loc.gov>)
 - LCNAF (Library of Congress Name Authority File) (<http://authorities.loc.gov>)
 - IPTC (International Press Telecommunications Council) (www.iptc.org)
 - Moving Image Genre-Form Guide (www.loc.gov/rr/mopic/migintro.html)
 - Taxonomy Warehouse <http://www.taxonomywarehouse.com/index.asp>)
- Review: Metadata standards for Cinematographic Works: http://filmstandards.org/fsc/index.php/Type_and_relationship_vocabularies#Vocabulary_initiatives

Topics/activities:

- Review taxonomy construction and controlled vocabulary standards: LCSH, LCNAF, Moving Image Genre-Form Guide
- How to create a data dictionary
- Controlled vocabulary exercise
- Logical data modeling: Entities, Relations, Attributes
 - Activity: Creating attributes for an entity relation model
- Database design – Physical modeling (Guest speaker: Deena Engel)
 - Database normalization (1st, 2nd, 3rd normal forms)
 - Look up tables
 - Join tables

Class 7: Mar. 28-- Metadata systems and tools; Databases Part 3 (Deena Engel)

Due this class:

- Reading: Churcher, Clare, *Beginning Database Design: From Novice to Professional*, Springer, 2012, Chapters 3 and 4. Available online through NYU Libraries.
- Reading: *Metadata*. Chapter 6: Metadata Services, p. 211-244
- Coyle, Karen and Baker, Tom. (2009). *Guidelines for Dublin Core Application Profiles*. Dublin Core Metadata Initiative. <http://dublincore.org/documents/profile-guidelines/>
- University of Washington Libraries Metadata Implementation Group. (2011). *Data Dictionaries (a.k.a. Schemas and Metadata Application Profiles or MAPS)*. <http://www.lib.washington.edu/msd/pubcat/mig/datadicts>

Topics/activities:

- Review database structures and tools for managing metadata
- Learn about designing application profiles
- Learn about methods for sharing metadata records
- FileMaker Pro set up and walk through (Guest speaker: Deena Engel)
 - Database and data model
 - Vocabularies
 - Layouts
 - Searching

Class 8: April 4 Metadata interoperability and crosswalks; XSLT Workshop

Due this class:

- Due: Physical model of database
- Reading: *Metadata*: p. 119-122; p. 267-284. (Available from Drive, see above)
- Reading: Chan, Lois Mai and Zeng, Marcia Lei, “Metadata Interoperability and Standardization: a Study of Methodology Part 1”, *Dlib Magazine*, June 2006
<http://www.dlib.org/dlib/june06/chan/06chan.html>
- Reading: Chan, Lois Mai and Zeng, Marcia Lei, “Metadata Interoperability and Standardization: a Study of Methodology Part 2”, *Dlib Magazine*, June 2006
<http://www.dlib.org/dlib/june06/zeng/06zeng.html>
- Reading: Churcher, Clare, *Beginning Database Design: From Novice to Professional*, Springer, 2012, Chapters 7 and 8. **Online version available through NYU Libraries.**

Topics/activities:

- Explore issues in metadata interoperability and crosswalks
- Consider crosswalks for data structure standards
- XSLT Workshop (Guest speaker: Ashley Blewer)

Class 9: April 11– Cataloging workshop

Guest speaker: Andrea Leigh, Head, Moving Image Processing, Library of Congress

Due this week:

- **ASSIGNMENT #2: Data Mapping exercise**
- *Descriptive Metadata for Television*. Pages 106-112.
- Review: <http://americanarchive.org/>

Topics/activities:

- Andrea will address cataloging moving images at the Library of Congress and will review use of PBCore.
- We will fully catalog a work together in class using different standards.
- Case study: American Archive of Public Broadcasting
- Cataloging levels and workflows

Class 10 (MAKE-UP): April 14 – Migrating and managing metadata. Guest speaker: Thelma Ross, MoMA Film Study Center

Topics/activities:

- Thelma will discuss and demonstrate the process of migrating existing metadata into a new database and structuring it according to a database model

- Metadata quality and dealing with messy data
- Exercise: cleaning up metadata (OpenRefine)

Class 11: April 18 -- Preservation, technical, and legal data

Due this class:

- Initial FMP database setup
- Reading: Caplan, Priscilla. *Understanding PREMIS*. Library of Congress, 2009. **[on NYU Classes]**
- Reading: *Descriptive Metadata for Television*. Pages 61-75
- Reading: Whalen, Maureen. "Rights Metadata Made Simple." In *Introduction to Metadata*. Online Edition, Version 3.0
http://www.getty.edu/research/conducting_research/standards/intrometadata/rights.html
- For the following, just get a general familiarity with them:
- Review : PREMIS 3.0 <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
- Review: SMPTE RP-210 (technical metadata dictionary) **[on NYU Classes]**

Topics/activities:

- Review data requirements and standards for technical, preservation, and legal metadata, including sources for controlled vocabularies for terms
- Discuss and compare the SMPTE Metadata Dictionary, PREMIS
- Review structuring legal data so the legal due diligence process is captured and reports can be generated
- Discuss data record construction, incorporating descriptive, physical, technical, legal, and preservation data
- Identifiers: p. 52-54; Metadata: p. 58-85
- Activity: Extracting technical metadata

Class 12: April 25—Managing and using metadata. Guest speaker: Seth Kaufman, Collective Access

Due this class:

- FMP Database design draft
- Review: *METS Primer*,
<http://www.loc.gov/standards/mets/METSPrimerRevised.pdf>
- Reading: National Information Standards Organization. *A Framework of Guidance for Building Good Digital Collections*. 3rd edition, December 2007,
<http://www.niso.org/publications/rp/framework3.pdf>
- Reading: *Descriptive Metadata for Television*. Pages 22-36. (Available from Drive, see above)
- Reading: *Audio-Visual Format Documentation Project: Background Paper*, Federal Agencies Digitization Guidelines Initiative, Audio-Visual Working Group
<http://www.digitizationguidelines.gov/guidelines/FADGI->

[AV AppSpecProj Bkgd 101007.pdf](#) (also review:
<http://www.digitizationguidelines.gov/audio-visual/>)

Topics/activities:

- Presentation on collective access and how it works (Guest speaker: Seth Kaufman)
- Containers and wrappers
- Metadata Encoding and Transmission Standard (METS)
- METS and extension schemas
- Using METS as a presentation and preservation format

Class 13: May 2 – Linked Data; Linked Data tools workshop

Due this class:

- FMP Database layout draft
- Reading: Coyle, Karen. “Library Data in a Modern Context”, in *Understanding the Semantic Web: Bibliographic Data and Metadata*, Library Technology Reports, January 2010 [on NYU Classes]
- Reading: “Library Linked Data Incubator Group Final Report” (25 Aug. 2011) <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>
- Reading: Van Maissen, Kara. *Bibframe AV Modeling Study: Defining a Flexible Model for Description of Audiovisual Resources*. (submitted May 15, 2014). <http://www.loc.gov/bibframe/pdf/bibframe-avmodelingstudy-may15-2014.pdf> (I encourage you to read it all, but at least Introduction (p. 2-9), parts 4 and 5 (p. 24-43))

Topics/activities:

- Review what Linked Data is and how it is beneficial to libraries, museums and archives
- Review Semantic Web relevant technologies
- Discuss use cases and how institutions are making their data available as LD
- Discuss the BIBFRAME Linked Data Model and how it is being adapted for audiovisual resources
- Workshop on Linked Data tools (Guest speaker: Matt Miller, NYPL)

Class 14: May 9 – Student Presentations

Due this class:

- **ASSIGNMENT #3: Final database with sample records**

Topics/activities:

- Student Presentations