MOVING IMAGE ARCHIVING & PRESERVATION PROGRAM
CINE-GT 1803: METADATA FOR MOVING IMAGE COLLECTIONS

This will be available on NYU Classes and updated periodically.
Spring 2017 – Mondays, 6 pm – 10 pm, 721 Broadway room 652
Instructor: Rebecca Guenther
NOTE: See in NYU Classes under Lessons for links to readings, assignments, etc.

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

Office hours: Mondays 4-6pm by appointment. 721 Broadway room 679
Class dates: Jan. 23; Jan. 30; Feb. 6; Feb. 27; Mar. 6; Mar. 20; Mar. 27; Apr. 3; Apr. 10; Apr. 17; Apr. 24; May 1; May 8
MAKEUP CLASS: Field trip to LCPA (Date to be determined)

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. 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, data quality, 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

ATTENDANCE:

Attendance at all classes is expected unless excused. More than one unexcused absence will affect grading.

GRADING:

Grades will be based on the following:
- Class participation and attendance (20%)
- Class activities/homework (10%);
- Data mapping project (30%)
- Final metadata project (40%)

Feedback on assignments will be given electronically.

TECHNOLOGY NEEDS:

Students are required to bring their own laptops to class each week. Both Windows and Mac are acceptable; pending they meet the following minimum requirements:

Mac
OS 10.6.8 or later
Intel Processor
At least 2 GB RAM
At least 30 GB available disk space

Windows
XP or later
Preferably 64 bit
At least 2 GB RAM
At least 30 GB available disk space

If you do not have access to a laptop, or do not have one that meets these minimum requirements, MIAP has some available for use in the class. Please see Ethan Gates or Kathy Scott.

Software needs
MIAP will provide licenses for the semester for the following:
- FileMaker Pro
- oXygen XML editor

IMPORTANT NYU TISCH POLICIES

Tisch Policy on Academic Integrity
The core of the educational experience at the Tisch School of the Arts is the creation of original work by students for the critical review of faculty members. Any attempt to evade that essential transaction through plagiarism or cheating is educationally self-defeating and a grave violation of Tisch's community standards. Plagiarism is presenting someone else's original work as if it were your own; cheating is an attempt to deceive a faculty member into believing that your mastery of a subject or discipline is greater than it really is. Penalties for violations of Tisch's Academic Integrity Policy may range from being required to redo an assignment to dismissal from the School. For more information on the policy--including academic integrity resources, investigation procedures, and penalties--please refer to the Policies and Procedures Handbook on the website of the Tisch Office of Student Affairs.

Non-Discrimination and Anti-Harassment Policy &
Sexual Misconduct, Relationship Violence, and Stalking Policy
Please read the Non-Discrimination and Anti-Harassment and Sexual Misconduct, Relationship Violence, and Stalking policies concerning equal treatment and opportunity for students and the maintenance of a safe living and learning environment that is free of bias, prejudice, discrimination, harassment, and violence. Harassment (e.g., name-calling, verbal insult, racial or sexual slurs) is not tolerated.

TEXTS: The following are texts used throughout the course, along with the articles listed below in the class descriptions. These are available on NYU Google Drive as indicated. They are also available on Course reserves at NYU Bobst Library.

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

- *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=0B15icbsejHfMcGJhc3ZteDNVXzg&usp=sharing

ASSIGNMENTS

File submission format for assignments: Ysemester_course number_author's last name_assignment number.file extension.

Example: 15s_1803_Smith_a1.doc

#1: In class assignments and homework. There will be several activities that we will do in class individually or in groups to complement the lectures. They may be completed in class or as homework. (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%)

#2: Final 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 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. Finally, you will populate the database with a set of 5 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. 6. Physical model draft due Mar. 6. Application profile and initial database setup due Apr. 17. Final assignment due May 8 (40%)

MIAP DIGITAL ARCHIVE: Your final projects 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. See above file-naming conventions.

Outline of topics, activities and assignments (makeup class: Trip to NYPL Performing Arts Library, date TBD)

	Date	Topic	Activities and assignments
1	1/23/17	Introduction to metadata	Activity: Introduction to metadata
2	1/30/17	Databases part 1; Metadata models	Workshop on databases part 1 Activity: Conceptual data modeling
3	2/6/17	Data structure standards; Databases part 2	Workshop on databases part 2 Due: Proposal for collection
4	2/27/17	Data structure standards; descriptive and archival standards; Databases part 3	Activity: comparison of EAD, MARC 21 and MODS Workshop on databases part 3 Due: Homework on databases
5	3/6/17	Presentation on ArchivesSpace; Controlled vocabularies and authority control	Sally Vermaaten: ArchivesSpace Activity: Controlled vocabularies Due this week: Physical model of final project draft
6	3/20/17	Metadata syntaxes; XML	Activity: Creating XML Homework: XML
7	3/27/17	Metadata interoperability; Application profiles	Activity: Creating application profiles Followup workshop on databases Due: XML Homework
8	4/3/17	Preservation, technical and legal data; Follow-up database workshop	Activity: Technical metadata

9	4/10/17	Cataloging workshop	Workshop by Andrea Leigh Due: Assignment #1: Data mapping
10	4/17/17	METS; Migrating and managing metadata	Due this week: Application profile and initial FMP database setup Workshop on database management: Thelma Ross
11	4/24/17	Linked Data	Workshop on Linked Data tools (Matt Miller)
12	5/1/17	Metadata systems and tools; Metadata quality and remediation	Presentation on Collective Access (Seth Kaufman) Activity: OpenRefine
13	Makeup TBD	Field trip to LPA	Explore metadata workflows, creation and management with Harrison Behl of the Rodgers & Hammerstein Archive of Recorded Sound
14	5/8/17	Student presentations	Due: Final assignment

Class 1: January 23– Introduction to Metadata

Due this class:

- Reading: Gilliland, Anne. (2008). "Setting the Stage," in *Introduction to Metadata*. http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html
- Reading: *Metadata.* Pages 3-22 Available on NYU Google Drive https://drive.google.com/a/nyu.edu/file/d/0B3lZH3Qmx14seUhQVGNJdGVLSEU/view?usp=sharing
- Reading: *Descriptive Metadata for Television*. Pages 1-18. Available on NYU Google Drive. https://drive.google.com/a/nyu.edu/file/d/0B15icbsejHfMZmJPNFZaNW5MazA/view?usp=sharing

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
- 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
- Activity: What is metadata and what is it used for?

Class 2: January 30 – Databases Part 1 (Deena Engel); 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**

- 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**
- 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/marcld-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:

- Database Workshop #1: 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)
    - Introduction to Filemaker Pro
    - Example of "Donahue" database
- Metadata models for library, archive and museum contexts
    - Why do we need metadata models?
    - FRBR (Functional Requirements for Bibliographic Records) and FRBR's applicability to moving image materials. http://www.ifla.org/VII/s13/frbr/frbr.htm
    - Introduction to Linked data model
    - BIBFRAME model
    - EN 15907 (***Film identification - Enhancing interoperability of metadata***
- Introduction to data models: Entities and Relations
    - Notation (e.g. crows foot, UML)
    - Creating an E-R model
- Homework: Conceptual Data modeling exercise

Class 3: Feb. 6 – Data structure standards and element sets (part 1: general); Databases Part 2 (Deena Engel)

Due this class:

- Proposed collection for final project (submit by email)
- Homework on conceptual data modeling
- 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**
- Reading: Churcher, Clare, *Beginning Database Design: From Novice to Professional,* Springer, 2012, Chapters 1 and 2. **Online version available through NYU Libraries:**
- **https://getit.library.nyu.edu/go/9399302**

Note that some of the following readings on Data structure standards will be covered in the Feb. 27 class.

- Reading: *Metadata* Pages 402-420. Available on NYU Google Drive.
- **Handouts:** See handouts on individual data standards and Data Standards List **from NYU Classes** (in Handouts folder under Resources/Readings)
- Reading: *Descriptive Metadata for Television.* Pages 37-54; 113-130 (sample records) (Available from Drive, see above)

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, PB Core, EBU Core, SMPTE DMS-1
- Review metadata records
- Data structure creation using spreadsheets
- Database workshop #2: Database design and modeling (Guest speaker: Deena Engel)
  - Database normalization (1st, 2nd, 3rd normal forms)
  - Creating 1st and 2nd normal forms
  - ITunes example in Excel
  - E-R diagram of ITunes example

*************NO CLASS Feb. 13 (in Culpeper)**************************
*************NO CLASS Feb. 20 President's day ************************

Class 4: Feb. 27 -- Data structure standards and element sets (part 2); Descriptive and archival content standards; Databases Part 3 (Deena Engel)

Due this class:

- Homework assignment on databases
- 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
- Reading: Churcher, Clare, *Beginning Database Design: From Novice to Professional,* Springer, 2012, Chapters 3 and 4. Available online through NYU Libraries. **https://getit.library.nyu.edu/go/9399302**
- 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. 445-456. Available from NYU Google Drive
- Reading: *FIAF Moving Image Cataloging Manual*. International Federation of Film Archives, 2016, especially p. 1-11. Available in NYU Classes and at: http://www.fiafnet.org/images/tinyUpload/E-Resources/Commission-And-PIP-Resources/CDC-resources/20160920%20Fiaf%20Manual-WEB.pdf

Topics/activities:

- Review moving image standards (not covered last week): PBCore, EBU Core, DMS-1

- 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, MARCXML and MODS
- FileMaker Pro set up and walk through (Guest speaker: Deena Engel)
    - Database and data model
    - Vocabularies
    - Layouts
    - Searching
- Activity: Database normalization using Filemaker

Class 5: Mar. 6 -- ArchivesSpace demo; Controlled vocabularies, authority control and subject analysis

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
- Review:
    - Library of Congress Genre/Form Terms for Archival Materials: Moving Image Genre-Form Terms. (Updated through 2015). http://olacinc.org/drupal/capc_files/GenreFormHeadingsList.pdf
    - International Press Telecommunications Council. IPTC Newscodes Concept. http://cv.iptc.org/newscodes/mediatopic
    - European Film Gateway. EFG vocabularies I: value lists and types for EFG dataelements. **[on NYU Classes]**

Topics/activities:

- Presentation on ArchivesSpace for managing archival collections (Sally Vermaaten, NYU Libraries)
- Review taxonomy construction and controlled vocabulary standards: LCSH, LCNAF, Moving Image Genre-Form Guide
- How to create a data dictionary
- Activity: Controlled vocabularies

********** NO CLASS MONDAY MARCH 13 (Spring Recess) *********

Class 6: Mar. 20 –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. 124-136. Available from NY Google Drive

- Reading: *Descriptive Metadata for Television.* Pages 76-88. (Available from Drive, see above)

Topics/activities:

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

Class 7: Mar. 27-- Metadata interoperability and crosswalks; Application profiles

Due this class:

- Homework: Creating XML records
- Review: Riley, Jenn. "Seeing Standards: a Visualization of the Metadata Universe" http://www.dlib.indiana.edu/~jenlrile/metadatamap/
- Reading: *Metadata:* p. 347-355. Available from NYU Google Drive
- Woodley, Mary S. (2008). "Crosswalks, Metadata Harvesting, Federated Searching, Metasearching: Using Metadata to Connect Users and Information" in *Introduction to Metadata* http://www.getty.edu/research/publications/electronic_publications/intrometadata/path.html
- 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 Especially: http://www.lib.washington.edu/msd/pubcat/mig/datadicts/pnwaudio http://www.lib.washington.edu/msd/pubcat/mig/datadicts/pnwmovingimage and http://www.lib.washington.edu/msd/pubcat/mig/datadicts/jackson

Topics/activities:

- Learn about use and design of application profiles
- How do different metadata standards work together? Explore issues about metadata interoperability
- Creating and understanding crosswalks
- XSLT
- XML and databases
- Activity: creating a metadata crosswalk

Class 8: April 3: Preservation, technical and legal metadata; Followup database workshop

Due this class:

- Reading: Caplan, Priscilla. *Understanding PREMIS.* Library of Congress, REVISION? 2009. URI
- Reading: Whalen, Maureen. "Rights Metadata Made Simple." In *Introduction to Metadata*.

http://www.getty.edu/research/publications/electronic_publications/intrometadata/rights.html

- For the following, 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
- 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
- Activity: Extracting technical metadata
- Followup on databases: bring your questions/issues (Deena Engel)

Class 9: April 10– Cataloging workshop
Guest speaker: Andrea Leigh, Head, Moving Image Processing, Library of Congress

Due this class:

- **ASSIGNMENT #1: Data Mapping exercise**
- *Descriptive Metadata for Television.* Pages 106-112.
- Review: http://americanarchive.org/

Due this week:

- **Physical model of database**
- **Application profile for database**

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: April 17 – METS: Migrating and managing metadata. Guest speaker: Thelma Ross, MoMA Film Study Center

Due this class:

- Due this week: Application profile draft and initial FMP database setup
- Reading: *METS Overview* http://www.loc.gov/standards/mets/METSOverview.v2.html
- 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 61-75.

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
- Container formats for packaging metadata and content
- Metadata Encoding and Transmission Standard (METS)
    - What is a METS document?

    - METS and extension schemas
    - Using METS as a presentation and preservation format
- Activity: Migrating metadata

MAKEUP CLASS: Field trip to LCPA (Date to be determined)

Class 11: April 24 – Linked Data; Linked Data tools workshop

Due this class:

- OCLC. "Linked Data for Libraries". YouTube https://www.youtube.com/watch?v=fWfEYcnk8Z8
- Reading: Schreur, Philip Evan. *The Academy Unbound: Linked Data as Revolution.* https://journals.ala.org/index.php/lrts/article/view/5073/6144
- 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 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 12: May 1—Metadata systems and tools; Metadata quality and remediation. Guest speaker: Seth Kaufman, Collective Access

Due this class:

- Reading: *Descriptive Metadata for Television.* Pages 22-36. Available from Drive, see above
- Reading: *Metadata*. Pages 251-258; 267-277

Topics/activities:

- Presentation on collective access and how it works (Guest speaker: Seth Kaufman)
- Review database structures and tools for managing metadata
- Learn about methods for sharing metadata records
- Metadata quality and dealing with messy data
- Exercise: cleaning up metadata (OpenRefine)

Class 14: May 8 – Student Presentations

Due this class:

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

Topics/activities:

- Student Presentations