Midwest Big Data Hackathon Iowa

**Mentor Name**: Amanda Xu

**Project Name**: Prototype BIBFRAME 2.0 Modeling for Library Info Spotlight “Opera Land”

**Outcomes:** Make library info spotlight of operas into linked data model using BIBFRAME 2.0 and extended vocabularies, visible to the web, and easily engaged with library users

**Project Description**:

Library linked data promises to meet libraries’ need for agility in content delivery and user engagement. This project will demonstrate the prototype of BIBFRAME 2.0 modeling for work, instance, item, agent, topic, etc. from the local bibliographic records in Alma and external data sources, representing library info spotlight of operas in Opera Land, a collection of opera books, videos, sound recordings, streaming media, etc. interwoven into user’s online experience using Alma, LC MARC to BIBFRAME Transformation Service, RDF Translator, RDF Validator, RDFa Validator, Apache Jena Fuseki Server, etc.

The initial data set consists of six library bibliographic data in linked data model using BIBFRAME 2.0, representing an opera titled "The Marriage of Figaro." The sample files are in RDF/XML, JSONLD, RDFA, and N3 formats. More info about BIBFRAME 2.0 can be found in the overview of the LC BIBFRAME 2.0 Model, available from http://www.loc.gov/bibframe/docs/bibframe2-model.html

**Implementation Plan**:

The goals of the project are to 1) retrieve, map and convert selected data sets for the music work of operas from the datasets listed for the [Hackathon](http://bigdata.uiowa.edu/Datasets.pdf) into BIBFRAME 2.0 and extended vocabularies, e.g. schema.org, 2) make the conversion process and tools to work together, 3) enrich the bibliographic data and grow the space of Opera Land by zoning consumed data sets into social media, books, news, wiki, movies, etc. on a website, 4) serve RDF data over HTTP through SPARQL queries, and 5) enable Google Custom Search.

**Evaluation Plan**:

The project at this phase will be evaluated based on the completion of each task listed in the implementation plan, and establish the web services to make linked library resources visible via the web.

**Equipment, Materials, and/or Supplies Needed**:

Computer, data storage, backup and archive, servers for networking, database, applications and web

**Staffing Requirement:**

1. Bachelor’s degree in computer sciences or related field or equivalent combination of education and experience;
2. 3 years’ minimum Web applications or systems development experience;
3. Proficiency with programming languages such as PHP, Ruby, Python, jQuery, HTML, JavaScript and CSS;
4. Knowledge of triple stores, SPARQL, and RDF;
5. Experience performing data transfers utilizing software library or language APIs; and
6. Working knowledge of Apache, Amazon Web Services, etc.

**Dates: Oct. 8 and 9**

**Useful Links:**

1. External Data Sets:
   1. DBPedia lookup: <https://github.com/dbpedia/lookup>
   2. MusicBrainz: <https://musicbrainz.org/doc/Development/XML_Web_Service/Version_2>

* Music CDs: https://musicbrainz.org/release/356bcf92-7260-4a95-b5b8-334df0872593
  1. IMSLP
     + Performance and Sheet Music: <http://imslp.org/wiki/Le_nozze_di_Figaro,_K.492_(Mozart,_Wolfgang_Amadeus)>
  2. IMDB:
     + Video: <http://www.imdb.com/title/tt0253259>
  3. Boston Univ: <http://link.bu.edu/portal/Le-Nozze-di-Figaro--The-marriage-of-Figaro--Die/UTfBiY8o1DY/>

1. SPARQL: <https://en.wikipedia.org/wiki/SPARQL>
2. FUSEKI: <https://jena.apache.org/documentation/serving_data/> and download site: https://jena.apache.org/download/
3. Opera Public website: <http://operaplanet.org/>
4. Webhost – Site Ground: <https://www.siteground.com/web-hosting.htm>
5. Amazon EC2 – Virtual Server Hosting: <https://aws.amazon.com/ec2/>
6. Ubuntu Server: <https://www.ubuntu.com/download/server>
7. Google Custom Search: https://developers.google.com/custom-search/docs/structured\_data