

# The Linked Cinema Project

Linked Open Data at the Anthology Film Archives

Olivia Meehan, Melissa Nuber, Jing Zhao  
Pratt Institute, INFO 670: Linked Open Data, Spring 2022  
Professor Cristina Pattuelli

---

## Organization Profile

# ANTHOLOGY FILM ARCHIVES

---

## Anthology Film Archives

- Archive and Theater in the East Village
  - Preserve and collect films and organize screenings and other public programming
- Founded in 1970
- Specializes in lesser-known (or unknown), avant-garde, experimental, and independent films.

<http://anthologyfilmarchives.org/>



---

## Statement of Need



---

### Problems

- Screening information pages are dead-ends
- No way to browse the collection or past events by particular film, director, etc.
- Uncontrolled vocabulary

### Needs

- Enhance the linterlinking of core entities
- Ensure the consistency of data description
- Cement the relationship between AFA data and others, make data more discoverable.

Still From "The Flicker", retrieved from  
[http://anthologyfilmarchives.org/uploads/stills/st\\_20090804\\_074\\_500x400.png](http://anthologyfilmarchives.org/uploads/stills/st_20090804_074_500x400.png)

---

# Environmental Scan


## Linked Open Data at Carnegie Hall


- CH's data emphasizes events
- Each event entity has its own URI and includes metadata related to date, venue, and participants


**CH-LOD: how?**


1. Understanding our data: what did we need to express?  
*(think: THINGS and RELATIONSHIPS)*

Events: key elements

Where (venues) 

When (date, time) 

Who (performers) 

What (works being performed) 

## The Harvard Film Archive Project

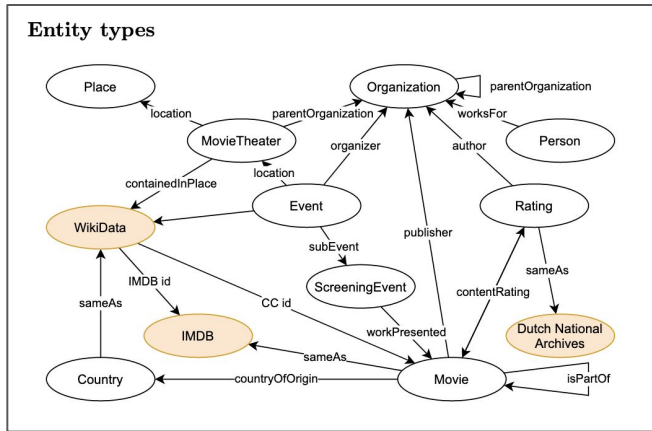
- Publish filmmaker data on Wikidata
- Ontology dictionary includes essential properties, suggested identifiers, and optional properties

Property	Value	Value may be new entity?	Usage note
archives at (P485)	institution holding the subject's archives	no	Harvard Film Archive's archival collections are listed here <sup>15</sup>
award received (P166)	film awards	no	
country of citizenship (P27)	country	no	
date of birth (P569)	year or more specific date	no	
date of death (P570)	year or more specific date	no	
employer (P108)	organization for which the subject works or worked	yes	
family name (P734)	surname	yes	

# Environmental Scan

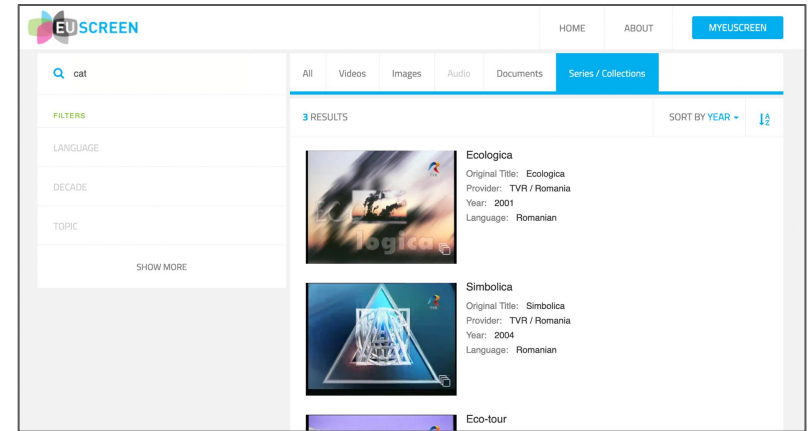
## Cinema Context

- Data model has 5 interconnected entity types
- The *schema.org* vocabulary was used. Dublin Core was supplemented in certain instances where *schema.org* fell short.



## The EUscreen Project

- Involved human-testing
- Keyword search with filters
- The front-end system can be hosted and moved to any location or multiple locations



# Project Design

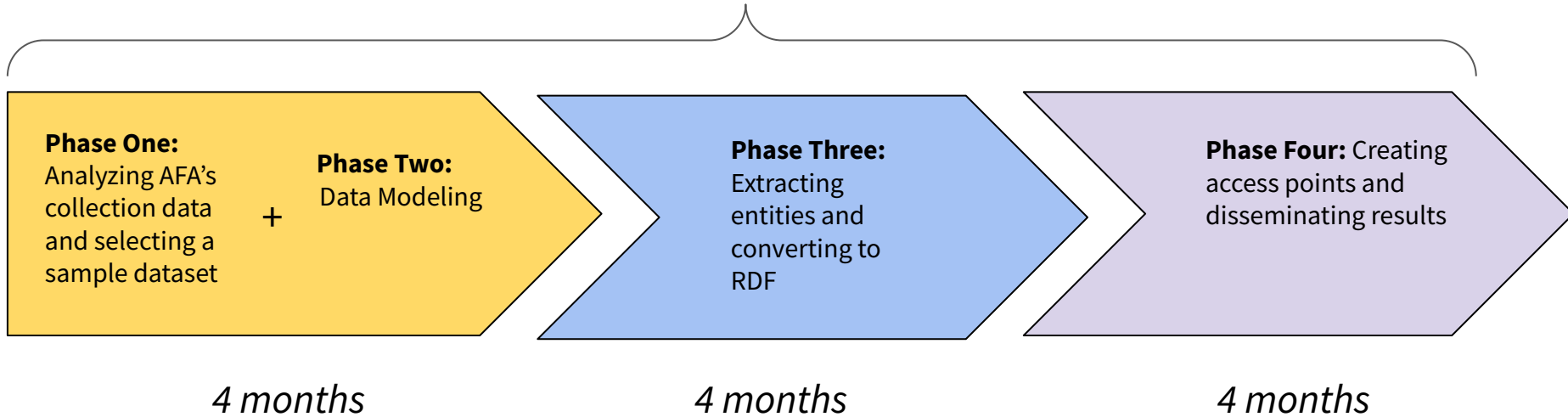
The Linked Cinema Project planning phase will be year a long project with four distinct, consecutive stages:

1. Analyzing AFA's collection data and selecting a sample dataset
2. Data modeling
3. Extracting entities and converting to RDF
4. Creating access points and disseminating results

**Wikibase** as hosting platform because it will allow us to use our own data model; manage our dataset openly; create our own entities and values; and channel our dataset to Wikidata.

# Proposed Project Timeline

Planning Phase: 12 months



# Project Personnel and Budget

**Personel-** The Linked Cinema Project will consist of **2** part time consultants

- **1** data scientists with Linked Open Data experience and
- **1** AFA archivist/subject specialist.

**Budget-** We are seeking **\$75,000**

- **\$60,000** will be allocated to cover project personnel
- **\$15,000** to cover equipment, technology and any logistical needs.

# Phase One: Inventory AFA and Select Sample Dataset

1. Perform a complete analysis of AFA's collection.
1. Select a sample of 90 items from AFA's repository that represent the our primary entities-
  - a. 30 **program** descriptions (digitized and available online)
  - b. 30 **film** notes (digitized and available online)
  - c. 30 **directors'** paper resources
1. Digitized resources converted to Microsoft Word documents and paper resources scanned and converted to plain to text.

**CITY SYMPHONIES**  
January 11 – March 10

This sprawling film series casts a wide net in order to bring together an enormous assortment of films that qualify as – or that in some way or other evoke – the genre of the City Symphony. The development of the City Symphony was closely intertwined with the dizzying pace of technological advancement and change that marked the advent of the 20<sup>th</sup> century, and the increased speed, capriciousness, and sense of fragmentation that this new era brought with it. The genre took shape during the 1920s, and became a particularly codified and dominant cinematic form into the following decade, a period in which dozens of variations on the genre were produced, including still-revered classics such as MANHATTAN (1921), RIEN QUE LES HEURES (1926), BERLIN, SYMPHONY OF A CITY (1927), MAN WITH A MOVIE CAMERA (1929), and many others. Emerging from the silent period and persisting into the first years of the sound era, these short and feature length works embodied the capabilities of the cinema at its most lyrical and dynamic. They combined the potential of the medium both as a tool for documenting cities, city-dwellers, and their changing environments, and as a more expressive, formally experimental means to capture the sensation of living in the city.

This series will showcase the seminal City Symphonies of the 1920s and 30s – the classics but also a generous selection of relatively obscure works – as well as those films produced up until today, and made throughout the world, that have continued the tradition of the City Symphony or that have found distinct ways of documenting and depicting cities and the experiences of those who inhabit them. For the purposes of the series, we're adopting a broad definition of the genre, to include films that – whatever their formal approach – strive not to convey the experience of a particular group of individuals within the city but to render the city a protagonist in and of itself.

\*City Symphonies\* is presented in conjunction with the publication of "The City Symphony Phenomenon: Cinema, Art, and Urban Modernity Between the Wars" (Routledge), edited by Steven Jacobs, Anthony Klink, and Eva Hietscher. Anthony Klink will be here in person for selected screenings. For more info regarding the book, visit: [www.routledge.com](http://www.routledge.com).

Example of program descriptions,  
[http://anthologyfilmarchives.org/film\\_screenings/series/50342](http://anthologyfilmarchives.org/film_screenings/series/50342)

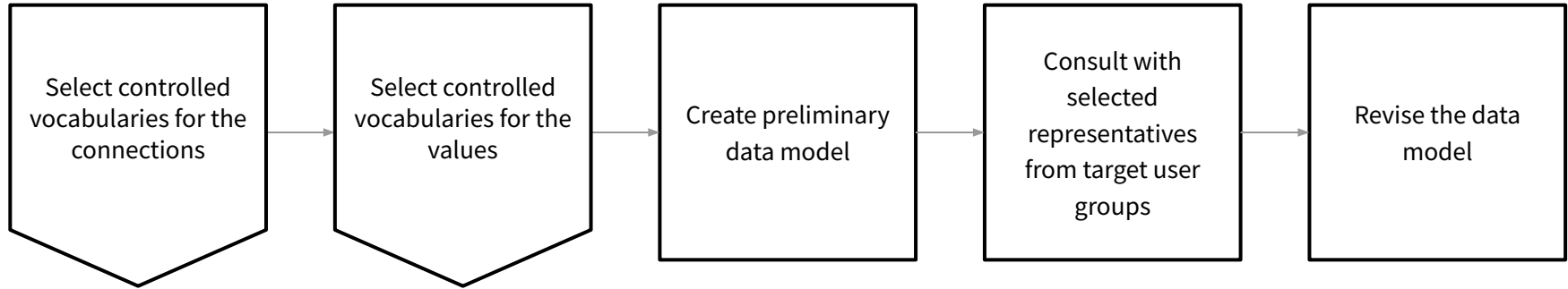
8:15 PM  
**CITY SYMPHONIES: MY WINNIPEG**  
by Guy Maddin  
2007, 80 min, 35mm  
**Share +**

This screening is part of: **CITY SYMPHONIES**

**Film Notes** ▲  
"Located at the intersection of documentary and dreamscape, Guy Maddin's brilliantly funny 'docu-fantasia' depicts his hometown – 'the heart of the heart' of North America, the official geographic center of the continent – as a vortex sucking up all of that landmass's wild, displaced emotional and sexual energy. Mixing the brazenly fictional with the incredibly stranger-than-fictional, MY WINNIPEG half-lies the honest truth." –TORONTO INTERNATIONAL FILM FESTIVAL

Example of film notes

# Phase Two: Data Modeling



WIKIDATA

Or

Linked.art  
CIDOC CRM  
schema.org

ULAN  
TGN  
AAT  
Library of Congress

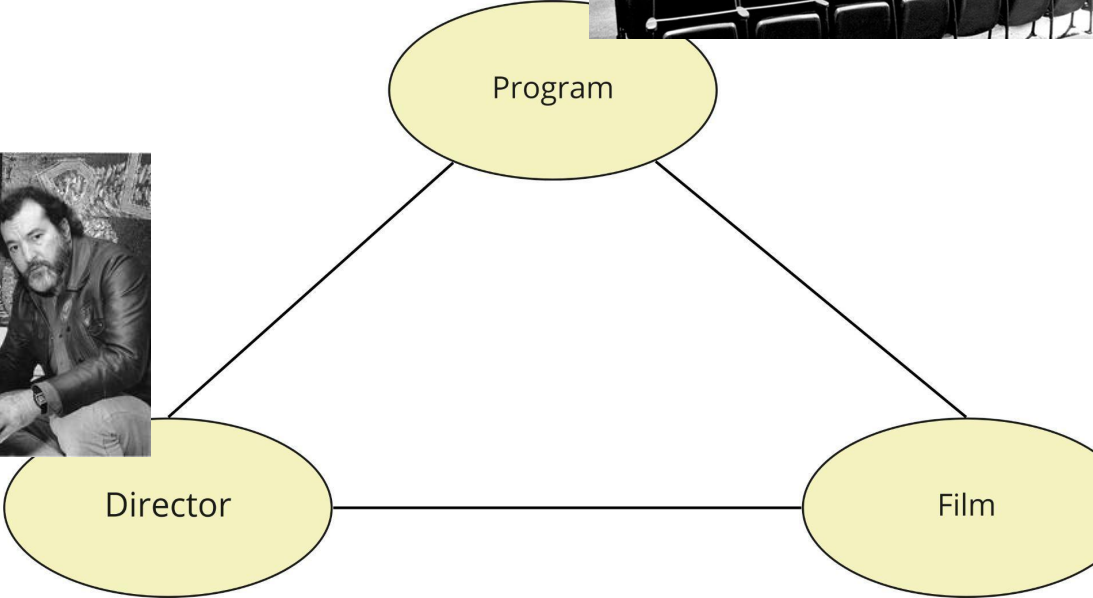
Or

Domain specific  
vocabularies

# Primary Entities



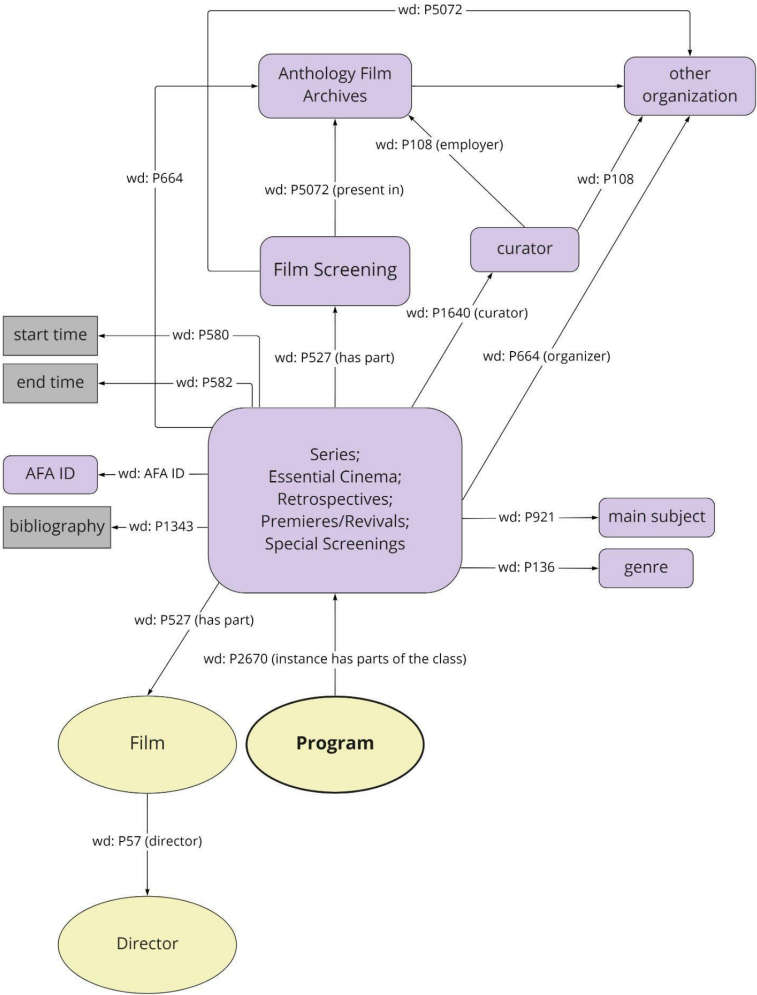
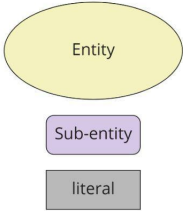
Photo of Paul Sharits, retrieved from <https://media.wnyc.org/i/800/0/h/85/1/Sharits.jpg>



Still From "T:O:U:C:H:I:N:G", retrieved from [http://anthologyfilmarchives.org/uploads/stills/st\\_20090827\\_072\\_500x400.png](http://anthologyfilmarchives.org/uploads/stills/st_20090827_072_500x400.png)

# Program

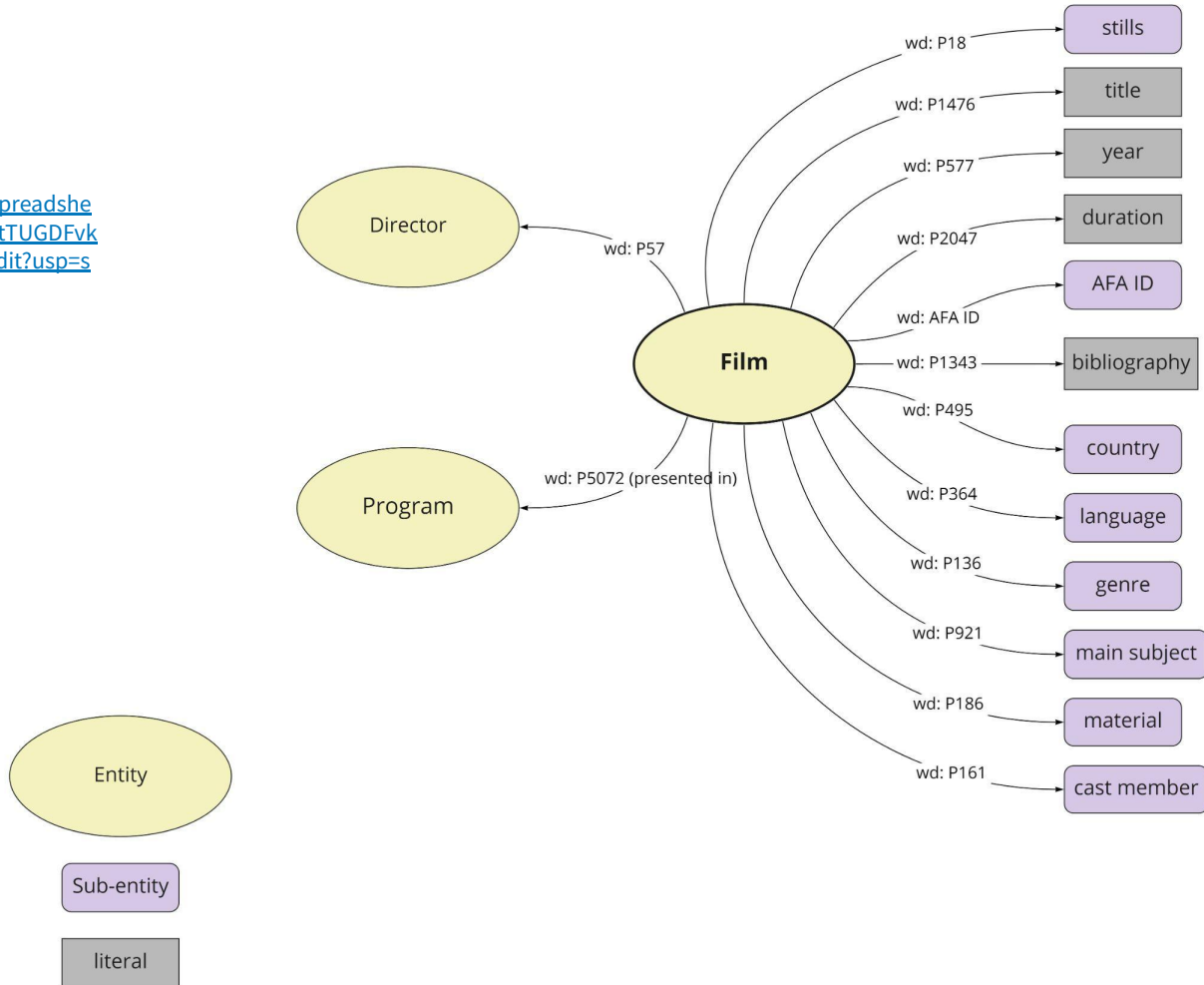
Spreadsheet:  
[https://docs.google.com/spreadsheets/d/1OHWFnlIDFnV9n3e9XINmq9KcnYvOOafxbeQYh\\_LopnE/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1OHWFnlIDFnV9n3e9XINmq9KcnYvOOafxbeQYh_LopnE/edit?usp=sharing)



# Film

Spreadsheet:

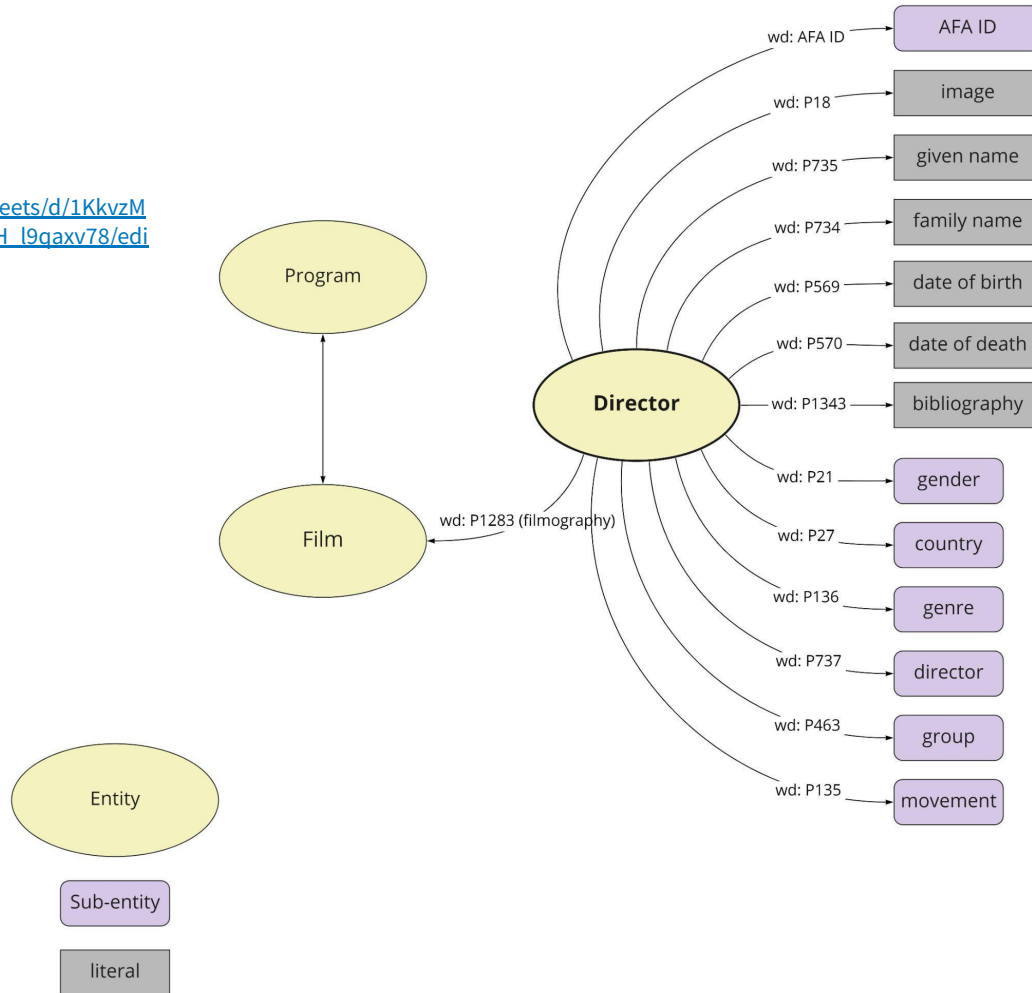
[https://docs.google.com/spreadsheets/d/1fM0ZwbSr1zHVAMztTUGDFvk\\_n3pfNPqz6Of0ku-3hwO/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1fM0ZwbSr1zHVAMztTUGDFvk_n3pfNPqz6Of0ku-3hwO/edit?usp=sharing)



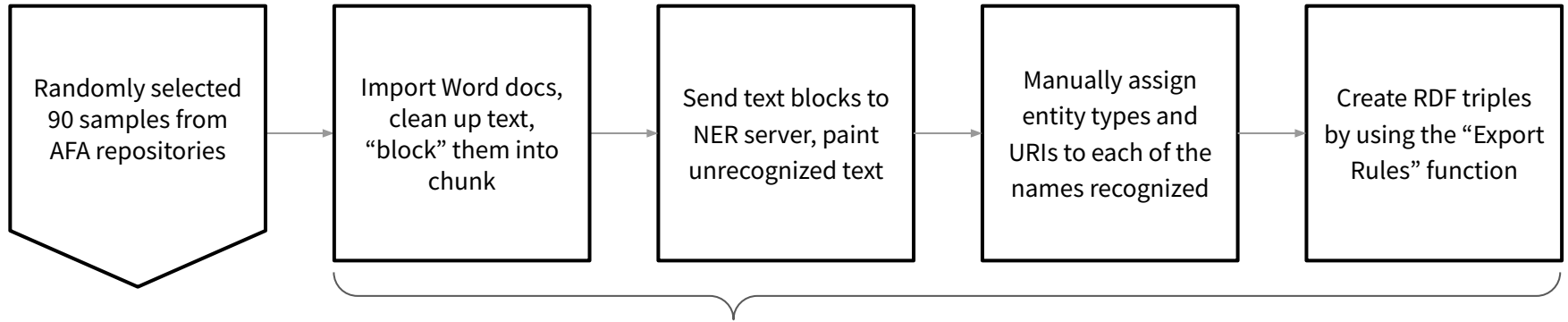
# Director

Spreadsheet:

[https://docs.google.com/spreadsheets/d/1KkvzM0UJ0MQIK3erFXBaLbk2-AAAny9T4IH\\_I9qaxv78/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1KkvzM0UJ0MQIK3erFXBaLbk2-AAAny9T4IH_I9qaxv78/edit?usp=sharing)



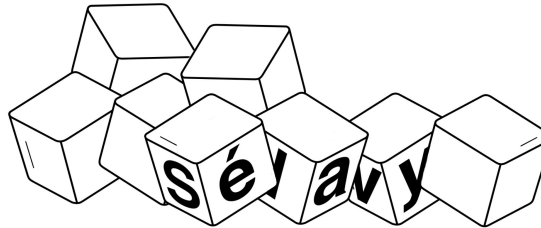
# Phase Three: Extract Entities and Create Triples



30 program descriptions

30 film notes

30 paper materials on individual directors



**Document to Data tool developed by the Semantic Lab at Pratt Institute**

Tool: <http://159.89.242.202:3000/>

Demo Video: [https://sem lab.io/howto/selavy\\_alpha](https://sem lab.io/howto/selavy_alpha)

# Phase Four: Creating Access Points and Disseminating Data

## Wikibase

- Export in JSON, XML, RDF etc.
- Perform SPARQL Queries

## Github

- Access and Collaboration with data

Propose New Wikidata Property (“Anthology Film Archives ID”)

Distribute project results at info shows/symposiums and create accompanying project video to present. Publish findings in open access repositories.



# Future Work

- Redesigning the front end of AFA's website to increase visibility of collections and improve search function.
- Researching target institutions and creating outreach and educational plans to encourage collaboration and linking to external datasets.
- Creating a task force to address archival/non-traditional film LOD standards.
- Planning an LOD symposium and inviting target institutions and partners to attend and participate.
- Designing a Linked Cinema Project website that aggregates linked archival materials from all partner institutions.
- Developing best practices for handling structured data related to nontraditional films including films with no known release dates, films of unique formats or durations, and orphan films where copyright holders are unknown.



**Thank you!**  
Questions?

# References

Carnegie Hall. “linked-data.” *GitHub*. Accessed April 27, 2022. <https://github.com/CarnegieHall/linked-data>.

Cinema Context. “About.” Accessed April 27, 2022.

<https://www.cinemacontext.nl/cgi/b/bib/bib-idx?c=cccfilm;sid=a50434085deae038e0e7d5f58f9c653;tpl=about.tpl;lang=en>.

Harvard Film Archive. “Cinematheque.” Accessed April 27, 2022. <https://harvardfilmarchive.org/cinematheque>.

Hudson, Rob. “Linked Open Data at Carnegie Hall.” February 8, 2022.

[https://docs.google.com/presentation/d/181WCQ3Rn2eQajGUXza\\_vVPx8lYKeEl-M-9QFTTYjNbo/edit](https://docs.google.com/presentation/d/181WCQ3Rn2eQajGUXza_vVPx8lYKeEl-M-9QFTTYjNbo/edit).

Museum and the Web. “Linking Europe’s Television Heritage.” Last modified April 14, 2012.

[https://www.museumsandtheweb.com/mw2012/papers/linking\\_europe\\_s\\_television\\_heritage](https://www.museumsandtheweb.com/mw2012/papers/linking_europe_s_television_heritage).

Wikidata. “WikiProject Linked Data for Production/Harvard Film Archive Project.” Accessed April 27, 2022.

[https://m.wikidata.org/wiki/Wikidata:WikiProject\\_Linked\\_Data\\_for\\_Production/Harvard\\_Film\\_Archive\\_Project#References](https://m.wikidata.org/wiki/Wikidata:WikiProject_Linked_Data_for_Production/Harvard_Film_Archive_Project#References).