

# Beyond TEI:

---

Metadata for  
Digital Humanities

Carolyn Hansen | Sean Crowe  
@dhsi17 | #beyondTEI

# Acknowledgements

**Christina Harlow**, Metadata Librarian, Cornell University

<https://github.com/cmh2166/c4IMDCpres>

**Bill McMillin**, Digital Metadata Librarian, University of Cincinnati

**Eira Tansey**, Digital Archivist & University Records Manager,  
University of Cincinnati

**James Van Mil**, Digital Analyst and Developer Librarian,  
University of Cincinnati

# A little housekeeping...

- Data, slides, and exercises for this workshop available at: **[https://github.com/crowesn/  
beyond-tei](https://github.com/crowesn/beyond-tei)**
- Twitter: #beyondTEI #dhsi2017

What are we doing  
this week?  
(Syllabus review)

# Why go beyond TEI?

- TEI is a great standard, but not all DH projects have text-encoding components
- This course will address how metadata impacts:
  - Transcription
  - Text mining and analysis
  - Database building
  - Crowdsourcing
  - Interoperability
  - Data visualization

# Metadata 101

- What is it?
- What should it look like?
- Top 5 metadata myths
- Types of metadata

# Metadata is structured data

More Formal Definition (& a really great resource):

<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

# You may hear metadata referred to as...

- tags
- search terms
- headings
- “data about data”

# Types of Metadata

- descriptive
- structural
- administrative
  - rights management
  - preservation

# Descriptive Metadata

- most intuitive form of metadata; you use it all the time!
- important for discovery
- describes the content of a resource
  - includes things like title, creator, subject, etc.

# Structural metadata

- used for complex objects and explains the relationship between them
- Example: digitized chapters that form a book will have structural metadata that explains how the parts (chapters) relate to the whole (book)
- structural metadata is very important for complex files such as datasets or files that require mediation

# Administrative Metadata

- Information used in management of the resource
- Example: When the resource was created, what file type is the resource, Who can access it?
- Two subgroups of Administrative Metadata are:
  - Rights management: deals with intellectual property rights
  - Preservation: information needed to archive and preserve a resource

# Metadata should be...

- useful
- consistent
- interoperable
- flexible
- shared
- preserved (sometimes)

# Metadata Myths

- you must use an “established” schema or controlled vocabulary
- you must be a librarian or metadata specialist to create metadata
- you must preserve all of your metadata
- metadata is either “right” or “wrong” — it’s really about standards, best practices, and curatorial decisions

# Seeing Standards:

## A Visualization of the Metadata Universe

Content: Jenn Riley

Design: Devin Becker

*Work funded by the Indiana University Libraries' White Professional Development Award*

Copyright 2009-2010 Jenn Riley



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 United States License  
<http://creativecommons.org/licenses/by-nc-sa/3.0/us/>.

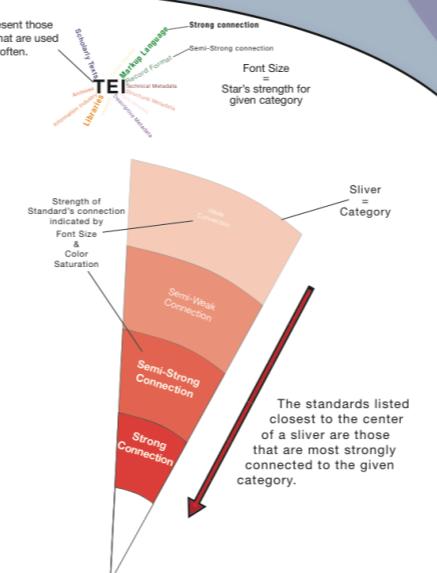
### LEGEND

#### Summary and Purpose

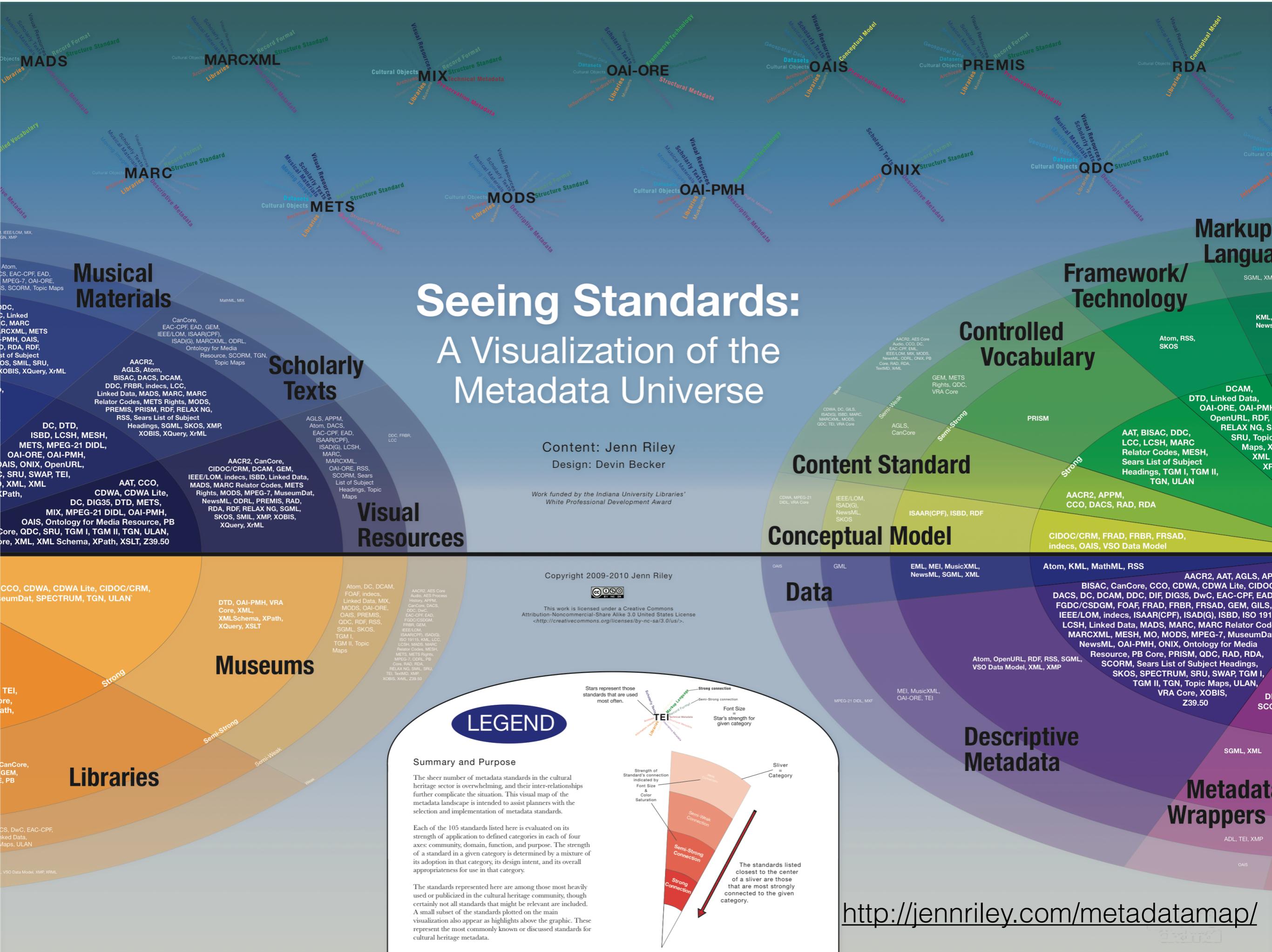
The sheer number of metadata standards in the cultural heritage sector is overwhelming, and their inter-relationships further complicate the situation. This visual map of the metadata landscape is intended to assist planners with the selection and implementation of metadata standards.

Each of the 105 standards listed here is evaluated on its strength of application to defined categories in each of four axes: community, domain, function, and purpose. The strength of a standard in a given category is determined by a mixture of its adoption in that category, its design intent, and its overall appropriateness for use in that category.

The standards represented here are among those most heavily used or publicized in the cultural heritage community, though certainly not all standards that might be relevant are included. A small subset of the standards plotted on the main visualization also appear as highlights above the graphic. These represent the most commonly known or discussed standards for cultural heritage metadata.



<http://jennriley.com/metadatauniverse/>



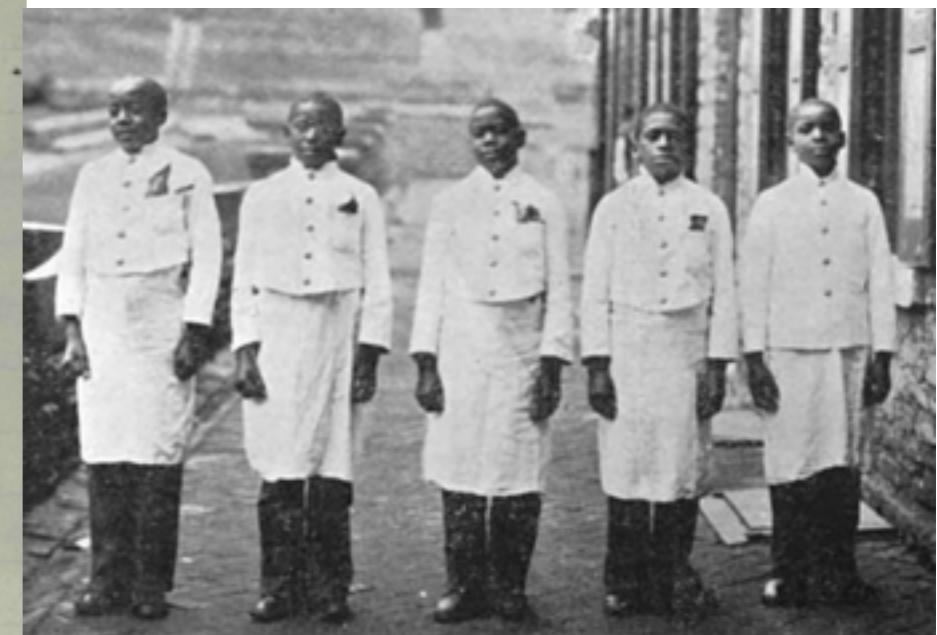
Exercise:  
Create Some Metadata  
About Another Student

(Custom Schema)

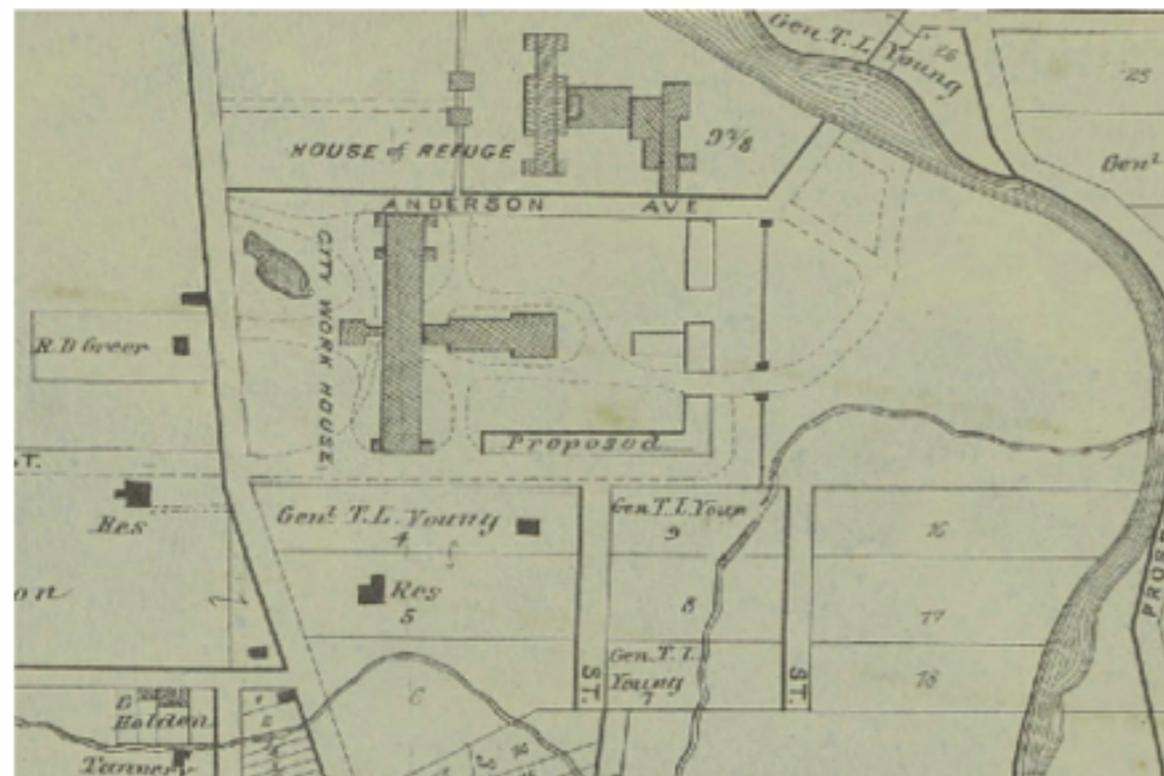
# The Cincinnati House of Refuge Dataset

- we'll be using this dataset for many of the exercises in this course; it contains over 6,000 admittance records
- data was transcribed from OCR'D original admittance ledgers from 1869-1902 held by the University of Cincinnati's Archives & Rare Books Library
- the ledgers contain rich demographic information about the children admitted, including crimes committed, parents' ethnicity/race, where the child was born, etc.
- presents interesting questions for standardization since historical data often contains contradictory information and outdated and/or offensive language

DATE	NO. OF COMMITMENT.	NAME	FOR WHAT COMMITTED.	BY WHOM COMMITTED.
1882 July 5	4691	William J. Sherlock	Incorrigibility	Police force House of Commons
	4692	Thomas Corbett	Breaking into cars	Police Court
	4693	William Corbett		
	44450	Andrew Freiherr	Incorrigibility	Recalled
	4270	John Higgins	Without suitable home	
	4694	William Martin	Incorrigibility	
	4695	John Maloney	Loitering	
	4696	Joseph Guttmiller		



Ledgers: <https://drc.libraries.uc.edu/handle/2374.UC/712586>



## Learn about Children, Crime, and Refuge in 19th century Cincinnati

From 1850 to 1912, thousands of poor and "criminal" children were sent to the Cincinnati House of Refuge for care and rehabilitation.

Explore what life was like for these children by searching our database of admission records, reading about the history of Refuge in Cincinnati, and viewing images, maps, and documents.



### Search Records

Search records of children's admissions to the House of Refuge from 1869-1902.

### The Refuge Story

Learn about the history of the House of Refuge and why children were admitted there.

### Browse Image Gallery

View images, maps, and documents about the House of Refuge.