Progress (and Pitfalls) of Linked Data Projects

Society of American Archivists 2018 Session 303
August 16, 2018
Linked (Open) Data: You’ve got questions...

Our speakers have answers.

1. Katherine Barbera, Assistant Archivist, Carnegie Mellon University, *What is Linked (Open) Data and Why Should You Care?*
2. Matthew R. Miguez, Metadata Librarian, Florida State University, *Rhizomes in the Archives: Linked Data and Mitigating Bias*
3. Samantha Norling, Digital Collections Manager, Newfields, *5-Star Linked Data: Maximizing the Power of LOD through Reconciliation*
4. Seth Shaw, Application Developer, University of Nevada, Las Vegas, University Libraries, *Linked Data Without (Necessarily) a Triplestore*
6. Toby Reiter, Web Developer, Archives of American Art, Smithsonian Institution, *Square Pegs in Round Holes: Creating Linked Open Data for Archives using the CIDOC CRM*
7. Elizabeth Ehrnst, Digital Initiatives Librarian, Georgia O’Keeffe Museum, *Cross Collection Linked Data Relationships*
8. Margaret Huang, Digital Archivist, Philadelphia Museum of Art, *Linked Data Across Collections, Institutions, and an Ocean: Duchamp Research Portal*
10. Greg Wiedeman, University Archivist, University at Albany, SUNY, *Challenges and Conflicts of Linked Data and Archives*
What is Linked (Open) Data?
And Why Should You Care?
What is L(O)D?

“Building Meaningful Connections”
What Is Linked (Open) Data? And Why Should You Care? SAA 2018
An Example: LOD on a Budget

Department of Film and Video Archive

- 18 Months
- 1 Unprocessed Collection (450 Linear Feet)
- 10,000 Digitized Objects
- 2 Project Archivists
- 7 Students & Volunteers
- 1 Project Developer
- 1 Database Manager
- 1 Webmaster
Why Linked Open Data?

Project Mottos:

1. Put Digital Audiences First.
2. Balance Access and Sustainability.
3. **Connections are Stories. Build Connections with Art.**
4. Don’t Repeat Yourself...or Others. And Share!
Carnegie Museum of Art
records.cmoa.org
Behind the Scenes

What Is Linked (Open) Data? And Why Should You Care? SAA 2018

Image Credit: David Newbury
Metadata

Events

- EventNumber
- EventID
- EventDate
- EventTitle
- EventNotes
- EventLinks

Catalogue

- CatalogueID
- CatalogueTitle
- CatalogueNotes
- CatalogueLinks

Parties

- PartyName
- PartyType
- PartyRole

Narratives

- Narratives
- NarrativeType
- NarrativeText

Multimedia

- MultimediaID
- MultimediaTitle
- MultimediaNotes

Rights

- Rights
- RightsType
- RightsConditions

69 Elements Per Object!
And Why Should You Care?

“Depth and Dimension”
Letter from Sally Dixon to Hollis Frampton (2/21/1972)

February 21, 1972

Hollis Frampton
River Road - nd #2
Hamilton, N.Y. 13346

Hollis, dear old one:

The Lorca "squib". I'm repeatedly reminded of B. Bailleuses sensibilities, especially the new "magician" film he's scripting now. Ask him, if you meet him, about it - rather, he will probably tell you.

How did the Elgin interlude fare? Having not heard a word, alternate between the assumptions that "no news is good news" and it was so disastrous that all are speechless. Please relieve my anxiety as soon as possible. Has Stan gone out? Back to Chicago or home? I haven't called Jane cause I'm so low on money, budget and personal.

I go to O'Grady's place (Buffalo) this coming weekend to recover from Van Buren's all night at Strasburgh Plantarium. (Victor G. wants to do something here and I'm hoping this (Rochester) will be a close enough precedent to sway our wooded headed local guardians of the stars, but I want to be able to speak from having seen with my own eyes. . . . !
Standish Lawder

Standish Dyer Lawder (1936 – 21 June 2014) was an American artist, art historian and inventor, who contributed to the structural film movement in the late 1960s and early 1970s.

— DBpedia

External Links

Wikipedia
Internet Movie Database

ARTWORKS IN THE PERMANENT COLLECTION

- Corridor, 1970, 1970
- Necrology, 1971, 1971
- Catfilm for Katy and Cynnie, 1972, 1972
A Few Years Later...

Dwindling Expertise

No Image Available

Letter from Sally Dixon to Robert Breer, 2/9/1970

Data Issues

Image Credit: © Robert Haller.
Thank You!

Kate Barbera
Assistant Archivist
kbarbera@andrew.cmu.edu
@brightarchives

Image Credit: © Robert Haller. fv001/001/004/001/B008/F08/001
Rhizomes in the Archives

Linked Data and Mitigating Bias
Representation & description

- Arrangement & description are not benign acts
- Conscious & unconscious biases influence description
- Hierarchical arrangements entrench power in the status-quo
  - Gilles Deleuze & Félix Guattari’s *A Thousand Plateaus*
The rhizome

- Data model conceived as an alternative to hierarchical structures
- Densely interconnected: many paths into, through, and out of
- Organic, serendipitous connection between resources
Graphs & networks

Mathematical structure composed of two elements:

- Nodes - points on a plane
- Edges - lines connecting those points

Linked data encodes a graph:

- Subjects and objects are nodes
- Predicates are edges
Graphs are rhizomes

Implementing linked data helps build the rhizome.
Why do we need rhizomes & linked data?

- Greater descriptive participation
  - More voices = greater & more robust description
- Connect *des fonds*
  - Dispersed collections can be reunited through the greater descriptive graph
- Improved discovery
  - Users can define their own paths through the rhizome, enabling more serendipity
- Archives potentially have the most to contribute
References


5-Star Linked Data
Maximizing the Power of Linked Data through Reconciliation

Samantha Norling
Digital Collections Manager,
Newfields
SAA Annual Meeting
August 16, 2018
What is 5-Star Linked Data?

Available on the web (any format); open license (OL)

Machine-readable, reusable (RE) structured data

Non-proprietary, open format (OF)

Use of URIs to identify things

Link your data to other data (LD)

https://5stardata.info/en/
WHY LINK TO OTHER DATA?
RECONCILIATION STRATEGIES

## Existing Tools: OpenRefine

Reconciliation Services - Wikidata

<table>
<thead>
<tr>
<th>All</th>
<th>ID</th>
<th>NAME_FULL</th>
<th>NAME_FIRST</th>
<th>NAME_MIDDLE</th>
<th>NAME_LAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>209</td>
<td>Bill Blass</td>
<td>Bill</td>
<td></td>
<td>Blass</td>
</tr>
<tr>
<td>159</td>
<td>911</td>
<td>Janet Payne Bowles</td>
<td>Janet</td>
<td>Payne</td>
<td>Bowles</td>
</tr>
<tr>
<td>180</td>
<td>1022</td>
<td>Edward Ruscha</td>
<td>Edward</td>
<td></td>
<td>Ruscha</td>
</tr>
<tr>
<td>415</td>
<td>2224</td>
<td>Cirrus Editions, Ltd.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>526</td>
<td>2715</td>
<td>Towle Silversmiths</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>653</td>
<td>3985</td>
<td>Garo Z. Antreasian</td>
<td>Garo</td>
<td>Zareh</td>
<td>Antreasian</td>
</tr>
<tr>
<td>1706</td>
<td>623960</td>
<td>Agnes Martin</td>
<td>Agnes</td>
<td></td>
<td>Martin</td>
</tr>
<tr>
<td>2109</td>
<td>710335</td>
<td>Sol LeWitt</td>
<td>Sol</td>
<td></td>
<td>LeWitt</td>
</tr>
</tbody>
</table>

http://openrefine.org/
Existing Tools: OpenRefine
Reconciliation Services - Wikidata

Reconcile each cell to an entity of one of these types:
- human
  - ID: Q5
- painting
  - ID: Q3305213
- watercolor painting
  - ID: Q18761202
- business
  - ID: Q4830453

Also use relevant details from other columns:

<table>
<thead>
<tr>
<th>Column</th>
<th>Include?</th>
<th>As Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME_FIRST</td>
<td>✓</td>
<td>given name</td>
</tr>
<tr>
<td>NAME_MIDDLE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME_LAST</td>
<td>✓</td>
<td>family name</td>
</tr>
<tr>
<td>NAME_SUFFIX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME_ORGANIZATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME_ORGANIZATION_ALT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAME_TAXONOMIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAMES_ALT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEX</td>
<td>✓</td>
<td>sex or gender</td>
</tr>
<tr>
<td>ACTOR_TYPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUMMARY</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Existing Tools: OpenRefine

Reconciliation Services - Wikidata

http://openrefine.org/

No Match (organization)

Suggested Match

Confident Match
Existing Tools: OpenRefine

Reconciliation Services - Wikidata

<table>
<thead>
<tr>
<th>NAME_FULL</th>
<th>Library of Congr</th>
<th>ULAN ID</th>
<th>Wikidata ID</th>
<th>place of birth</th>
<th>place of birth coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill Blass</td>
<td>n86850097</td>
<td>500105667</td>
<td>Q861953</td>
<td>Fort Wayne</td>
<td>41.08045,-85.13915</td>
</tr>
<tr>
<td>Janet Payne Bowles</td>
<td>n93102397</td>
<td>500055908</td>
<td>Q19974756</td>
<td>Indianapolis</td>
<td>39.7677777777778,-86.15805555556</td>
</tr>
<tr>
<td>Edward Ruscha</td>
<td>n50022260</td>
<td>500024159</td>
<td>Q430967</td>
<td>Omaha</td>
<td>41.26666666666667,-96.01666666667</td>
</tr>
<tr>
<td>Cirrus Editions, Ltd.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Towle Silversmiths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garo Antreasian</td>
<td>n88614441</td>
<td>500111055</td>
<td>Q26155343</td>
<td>Indianapolis</td>
<td>39.7677777777778,-86.15805555556</td>
</tr>
<tr>
<td>Agnes Martin</td>
<td>n50040906</td>
<td>500024489</td>
<td>Q275610</td>
<td>Macklin</td>
<td>52.3333,-109.95</td>
</tr>
<tr>
<td>Sol LeWitt</td>
<td>n79092583</td>
<td>500115429</td>
<td>Q168587</td>
<td>Hartford</td>
<td>41.76333333333333,-72.685</td>
</tr>
</tbody>
</table>

http://openrefine.org/
### Existing Tools: OpenRefine

Fetch URL against SPARQL Endpoint or API

<table>
<thead>
<tr>
<th>OBJECT_TYPES</th>
<th>Getty SPARQL Query</th>
<th>Getty Query Return JSON</th>
<th>Getty Term</th>
<th>Getty URI</th>
</tr>
</thead>
</table>

But the best strategy for reconciliation is…

Incorporate linked data in ongoing cataloging activities!

- Does your archival management system (e.g., ArchivesSpace) support inclusion of URIs for controlled vocabularies?
- Add URIs to existing authority records
- Develop workflows that include reconciliation or linked data research, and train all cataloging staff
Reconciliation Resources:

Sami Norling
@SamiNorling
snorling@discovernewfields.org
Linked Data without (necessarily) a Triplestore

Seth Shaw | SAA 2018 | Session 303
Linked Data → RDF → Triplestore

Not Necessarily!
5 ★ OPEN DATA

★ make your stuff available on the Web under an open license
★★ make it available as structured data
★★★ make it available in a non-proprietary open format
★★★★ use URIs to denote things, so that people can point at your stuff
★★★★★ link your data to other data to provide context

So, if not RDF and triplestores, then what?
One option:

1. Meets the linked data requirements
   - Structured open format
   - URIs
2. A lot of code libraries make it easy to work with JSON
3. Store in a JSON-aware database or generate it on the fly
4. Search engines (e.g. Google) understand it
You may already be using it!

No triplestore necessary!
But I *want* a triplestore!

*Go ahead and populate one.*

E.g. Scrapy

E.g. Blazegraph
Coming soon... Islandora CLAW!

For all your SPARQL needs.
Photograph of the Flamingo Hotel at sunset circa mid 1950s
Go try it yourself!

Thank you.

Links
https://json-ld.org
https://archivesspace.org
https://scrapy.org
https://github.com/RDFLib
https://www.blazegraph.com
https://islandora.ca/CLAW

Image/Icon Sources
5 Star Steps: http://5stardata.info/images/5-star-steps.png
JSON-LD: https://json-ld.org/images/json-ld-logo.svg
Spider icon from The Noun Project by Edward Boatman: https://commons.wikimedia.org/wiki/File:Noun_project_175.svg
Cog wheels: https://commons.wikimedia.org/wiki/File:Cog-icon-grey.svg
RDF: http://www.dotnetrdf.org/
Islandora CLAW: http://islandora-claw.github.io/CLAW/
Drupal 8: https://www.drupal.org/about/media-kit/logos
A Minimum Viable Strategy for Archives and Linked Data using Schema.org

Mark A. Matienzo, Stanford University / @anarchivist
Society of American Archivists Session 303 #s303
16 August 2018
Rationale and objectives

- Group of archivists and technologists interested in pragmatic approaches to linked data about archives
- Investigating use of Schema.org and its extensions as a minimally viable mechanism for publishing linked data
- Development of mappings from archival description standards to Schema.org/extensions
- Production of RDF-modeled archival description directly from archives management systems
Schema.org

- Created in 2011 by Bing, Google, and Yahoo to address structured data format proliferation for search engines
- Provides single schema across range of topics: people, creative works, places, etc. (589 types, 862 properties)
- Used on 1+ billion web pages & many popular websites
- Expressible as JSON-LD, RDFa, and Microdata
- Provides extension mechanism for both Schema.org-hosted and external extensions
Schema Architypes

- Schema.org extension to represent archives, proposed for inclusion in October 2017
- Introduces a minimal set of new types/properties
- Selected as the basis for our data modeling work
Mapping description to Schema.org

- Preliminary mappings from ISAD(G), ISAAR-CPF, DACS, and ArchivesSpace/Atom data models
- Mostly straightforward with notable exceptions:
  - Description control
  - Level of description
  - Reference code
  - Precision of note types
Example

{
  "url": "http://archives.library.rice.edu/repositories/2/resources/1038",
  "@context": "http://schema.org/",
  "@type": ["Collection", "ArchiveComponent"],
  "name": "Houston Folk Music collection",
  "inLanguage": "EN",
  "holdingArchive": {
    "@type": ["Archive", "LocalBusiness"],
    "name": "Woodson Research Center Special Collections & Archives",
    "address": "Fondren Library, Rice University, 6100 Main, Houston, TX 77005",
    "url": "http://library.rice.edu/woodson/"
  },
  "creator": [
    { "@type": "person", "name": "Townes Van Zandt", "@id": "http://viaf.org/viaf/56876870" },
    { "@type": "person", "name": "Guy Clark", "@id": "http://viaf.org/viaf/29717590" },
    { "@type": "person", "name": "Vince Bell", "@id": "http://viaf.org/viaf/78141250" },
    { "@type": "person", "name": "Lyle Lovett", "@id": "http://viaf.org/viaf/120632356" }
  ],
  "description": ["This collection contains items documenting the Houston Folk music scene from the 1970s-1980s.",
    "During the 1960s-1980s, Houston, TX had a vibrant folk scene. Local musicians and those from other parts of Texas and the
    U.S. socialized and played at a variety of music establishments around the city, including Anderson Fair Retail Restaurant, Sand
    Mountain Coffeehouse, Liberty Hall, The Old Quarter, Theodore's and Corky's. This scene spawned many musicians, Townes Van
    Zandt, Guy Clark, Robert Earl Keen, Nanci Griffith, Lynn Langham, Wrecks Bell, Eric Taylor, John Vandiver, Danny Everitt, Vince
    Bell, Richard Dobson, Don Sanders, Wheatfield/St. Elmo's Fire, Dogtooth Violet, Bill Cade, Lyle Lovett, and many more."
  ],
  "dateCreated": "1975-1985",
  "accessConditions": ["This material is open for research.",
    "Permission to publish from this material must be obtained from the specific copyright owner."
  ],
  "ownershipInfo": "Albums donated by Brendan Doss, III. Townes Van Zandt reel donated by Craig Calvert."
}
Expected benefits and future work

- Mappings allow for publication of linked data directly from archives management systems or discovery environments
- Development of a recommended profile for implementation in a wide variety of systems
- Further investigation and mapping to other ontologies and data models, and addressing existing gaps in modeling
Thank You!

Resources

- Archives and Linked Data Interest Group: [https://archival.github.io/](https://archival.github.io/)
- W3C Schema Architypes Community Group: [https://www.w3.org/community/architypes/](https://www.w3.org/community/architypes/)

Acknowledgements

The Archives and Linked Data Interest Group:

- Scott Carlson (Rice University)
- Mark Custer (Yale University)
- Patrick Galligan (Rockefeller Archives Center)
- Dan Gillean (Artefactual Systems)
- Gloria Gonzalez (Zepheira)
- Maggie Hughes (UCLA)
- Mark Matienzo (Stanford University)
- Dave Mayo (Harvard University)
- Laney McGlohon (LYRASIS/ArchivesSpace)
- Evelyn McLellan (Artefactual Systems)
- Katy Rawdon (Temple University)
- Elizabeth Russey Roke (Emory University)
- Ruth Kitchin Tillman (Penn State)
Square Pegs in Round Holes

Creating Linked Open Data for Archives using the CIDOC/CRM

Toby Reiter
Archives of American Art
• A consortium project to create a shared resource for research and discovery in American Art

• The Collaborative currently includes 14 partner institutions (full list at http://americanartcollaborative.org)
Concept

- Create a common data source on American Art
- Published as LOD
- Using linked.art, a subset of the CIDOC-CRM (a reference model for Cultural Heritage Documentation)
- A good fit for museum objects and the AAC
Complications

• Well…for most members of the AAC
Complications

• Well…for most members of the AAC
  • Archival materials ≠ museum objects
  • Archival description ≠ object-level description
• Context is key
• Finding aids?
Complications (cont.)

• Finding Aids? Not really.
Complications (cont.)

- Finding Aids? Not really.
- Archival collections not necessarily supported in the model
- Contributed collection data, but only titles, dates, creator information
Complications (there’s more!)

E.g. Collection Creator

**MARC:** 100 $a

**CIDOC/CRM:** E78_Collection -> p108i_was_produced_by -> E12_Production -> p14_was_carried_out_by -> E39_Actor
Complications (there’s more!)

E.g. Collection Creator

**MARC:** 100 $a

**CIDOC/CRM:** E78_Collection -> p108i_was_produced_by -> E12_Production -> p14_was_carried_out_by -> E39_Actor

**BibFrame:** Work -> creator -> Agent
Complications (update)

- Finding Aids? Maybe!
- EAD mapping to CIDOC/CRM
- http://www.cidoc-crm.org/Resources/the-ead

Toby Reiter
reitert@si.edu
Archives of American Art
https://www.aaa.si.edu
http://data.aaa.si.edu

Square Pegs in Round Holes
Creating Linked Open Data for Archives using the CIDOC/CRM
SAA 2018 – Session 303
“To create one’s own world takes courage.” - Georgia O’Keeffe
DigIn project overview

2016 | Data Management Working Group formed
- Goal: Improve the core digital infrastructure, establishing standards and data management practices that serve interoperability
- Staff from archives, curatorial, collections management (fine art, personal and historic property), and library

2017-2018 | IMLS Museums for America Grant received
- Goal: Through a unified digital interface, expand access to the breadth and depth of the Georgia O'Keeffe Museum holdings and facilitate the discovery of new relationships
- Building upon the work of the American Art Collaborative (AAC)
- Developer: Design for Context

Georgia O'Keeffe Museum
Types of data to be published

- Authority files
- Bibliographic data
- Digital images
- Encoded archival descriptions
- Geographic data
- Museum object data

Georgia O’Keeffe Museum
Data models

CIDOC Conceptual Reference Model (CRM)
- Provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation [cidoc-crm.org](http://cidoc-crm.org)

Linked Art
- Community working together to create a shared model based on Linked Open Data to describe Art [linked.art](http://linked.art)
- CIDOC-CRM as the core ontology
- Getty Vocabularies (AAT, ULAN, TGN) as core sources of identity
- JSON-LD as the primary target serialization
Archives data model

Model selection based on an ontology for the archives that could be easily aligned with Linked Art and could be part of an integrated environment

Part of the archives model at collection level
Collection level data in Pubby, a linked data frontend for SPARQL endpoints

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>crm:P104_is_subject_to</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/accessrestriction/aspect2556c6a93464a252565caee77c7">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/accessrestriction/aspect2556c6a93464a252565caee77c7</a></td>
</tr>
<tr>
<td>crm:P108I_was_produced_by</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/production">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/production</a></td>
</tr>
<tr>
<td>crm:P1_is_identified_by</td>
<td>[1: anonymous resource]</td>
</tr>
<tr>
<td>crm:P2_has_type</td>
<td><a href="http://vocab.getty.edu/aat/300375748">http://vocab.getty.edu/aat/300375748</a></td>
</tr>
<tr>
<td>crm:P2_has_type</td>
<td><a href="http://vocab.getty.edu/aat/collection">http://vocab.getty.edu/aat/collection</a></td>
</tr>
<tr>
<td>crm:P30I_custody_transferred_through</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/acquisition/aspect24365424951b3e35">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/acquisition/aspect24365424951b3e35</a></td>
</tr>
<tr>
<td>crm:P33_used_specific_technique</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/processinfo/aspect26dcd5201b265e1a23760b081b743f7b21">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/processinfo/aspect26dcd5201b265e1a23760b081b743f7b21</a></td>
</tr>
<tr>
<td>crm:P43_has_dimension</td>
<td>[1: anonymous resource]</td>
</tr>
<tr>
<td>crm:P45_consists_of</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform260269b7">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform260269b7</a></td>
</tr>
<tr>
<td>crm:P45_consists_of</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform23857861">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform23857861</a></td>
</tr>
<tr>
<td>crm:P45_consists_of</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform24576383">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/controlaccess/genreform24576383</a></td>
</tr>
<tr>
<td>crm:P67I_refers_to</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422</a></td>
</tr>
<tr>
<td>crm:P67I_refers_to</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422</a></td>
</tr>
<tr>
<td>crm:P67I_refers_to</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s36137145">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s36137145</a></td>
</tr>
<tr>
<td>crm:P67I_is_referred_to_by</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s2363422</a></td>
</tr>
<tr>
<td>crm:P67I_is_referred_to_by</td>
<td><a href="http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s36137145">http://172.20.0.10:8080/archive/archdesc/collection/PapersfromtheGhostRanchLibrary/biographis/s36137145</a></td>
</tr>
<tr>
<td>rdf:type</td>
<td>crm:E19_Physical_Object</td>
</tr>
</tbody>
</table>

This page shows information obtained from the SPARQL endpoint at http://localhost:query. As Turtle | As RDF/XML | Browse in Disco | Browse in Tabulator | Browse in OpenLink Browser | Georgia O'Keeffe Museum
Next steps

Still thinking about...

- More data
- More URIs
- Modeling of additional data
- Sustainability
- Tools for public access and visualization
Next steps

Team data visualization based on Nadieh Bremer’s Dragon Ball Z

Georgia O’Keeffe Museum
Linked data across collections, institutions, and an ocean: Duchamp Research Portal

Margaret (Marge) Huang, Project Manager
Digital Archivist
Philadelphia Museum of Art
Project Overview

Unify the Philadelphia Museum of Art’s digitized holdings with our partners, the Centre Pompidou and Association Marcel Duchamp – making the bulk of Duchamp’s archival materials, consisting of approximately 60,516 documents, accessible and discoverable through a single interface built on Linked Open Data and IIIF.
Project Background

• 2014-2016: Received NEH Humanities Collections and Reference Resources planning grant
  • Established a Duchamp Research Portal Advisory Board
  • Project meetings in Paris
  • Surveyed 15 collections
  • Circulated user survey to 300+ people with a nearly 50% return rate
  • Produced a summative white paper

• 2017: Received NEH Implementation grant
  • Kick off meetings in Paris
  • Began systematic digitization
  • Hired developers – Design for Context
  • Gathered sample data from Partners
Project Progress

- 2018: Information and Technical Infrastructure
  - Finished digitization
  - Sample data analysis – gaps and matches
  - Metadata survey – target data model
  - Technical specifications
  - Metadata creation
  - Goal: wireframes/proto website for December partners meeting
## Archives Data Fields - Name of Creator and/or other Personal Names

<table>
<thead>
<tr>
<th>Name of Creator and/or other Personal Names (CP = controlaccess:persname/ PMA = origination&amp;controlaccess:persname / AMD = auteur&amp;destinataire&amp;personnages&amp;photographe)</th>
<th>Minimum required field</th>
<th>Link to External Authority</th>
<th>Translate data</th>
</tr>
</thead>
</table>

## Archives - Comments about Name of Creator and/or other Personal Names

Your answer

## Archives Data Fields - Personal Name roles

<table>
<thead>
<tr>
<th>Personal Name roles (CP = role attribute / PMA = role attribute / AMD = auteur&amp;destinataire&amp;personnages&amp;photographe)</th>
<th>Minimum required field</th>
<th>Link to External Authority</th>
<th>Translate data</th>
</tr>
</thead>
</table>

## Archives - Comments about Personal Name Roles

Your answer

## Archives Data Fields - Geographic Name

| Geographic Name | Minimum required field | Link to External Authority | Translate data |
Target Data Models
# Target Data Models

<table>
<thead>
<tr>
<th>DACS Concept Area</th>
<th>Target Data Field</th>
<th>Label</th>
<th>Level</th>
<th>EAD Field</th>
<th>Attributes Required (in Addition to Element Values)</th>
<th>Required/Optional</th>
<th>Vocabulary of External Authority</th>
<th>Translate</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Elements</td>
<td>Language and Scripts of the Medium</td>
<td>Language</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>common_language</td>
<td>common_language</td>
<td>langcode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Elements</td>
<td>Physical Access</td>
<td>Physical</td>
<td></td>
<td>Location</td>
<td>&lt; xmlns=&quot;<a href="http://www.iso.org/iso/15500-1">http://www.iso.org/iso/15500-1</a>&quot; physloc id=&quot;head&quot; &gt;</td>
<td>required</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Elements</td>
<td>Conditions Governing Reproduction and Use</td>
<td>Rights Information</td>
<td>Object/Item</td>
<td>Folder</td>
<td>rights</td>
<td>rights</td>
<td>required</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Access Elements</td>
<td>Conditions Governing Reproduction and Use</td>
<td>Rights Information</td>
<td>Object/Item</td>
<td>Folder</td>
<td>rights</td>
<td>rights</td>
<td>required</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Access Elements</td>
<td>Immediate Source of Acquisition</td>
<td>Acquisition Information</td>
<td>Series/Sub-serial</td>
<td>Folder</td>
<td>acquisitioninfo</td>
<td>acquisitioninfo</td>
<td>preferred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Elements</td>
<td>Immediate Source of Acquisition</td>
<td>Acquisition Information</td>
<td>Series/Sub-serial</td>
<td>Folder</td>
<td>acquisitioninfo</td>
<td>acquisitioninfo</td>
<td>preferred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archival Description</td>
<td>Level of Description</td>
<td>Level of</td>
<td>Description</td>
<td>Object/Item</td>
<td>level</td>
<td>level</td>
<td>required</td>
<td>level = class, series, folder, file, fonds, item, other, level, subfolder, subseries</td>
<td></td>
</tr>
<tr>
<td>Archival Description</td>
<td>Level of Description</td>
<td>Level of</td>
<td>Description</td>
<td>Object/Item</td>
<td>level</td>
<td>level</td>
<td>required</td>
<td>level = class, series, folder, file, fonds, item, other, level, subfolder, subseries</td>
<td></td>
</tr>
<tr>
<td>Content and Structure Elements</td>
<td>Scope and Content Description</td>
<td>Collection/</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>scopecontentdesc</td>
<td>scopecontentdesc</td>
<td>preferred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content and Structure Elements</td>
<td>Scope and Content Description</td>
<td>Collection/</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>scopecontentdesc</td>
<td>scopecontentdesc</td>
<td>preferred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content and Structure Elements</td>
<td>Scope and Content Description</td>
<td>Object/Item</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>scopecontentdesc</td>
<td>scopecontentdesc</td>
<td>required</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content and Structure Elements</td>
<td>Scope and Content Description</td>
<td>Object/Item</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>scopecontentdesc</td>
<td>scopecontentdesc</td>
<td>required</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlled Access Headings</td>
<td>Location</td>
<td>Collection/</td>
<td>Sub-serial</td>
<td>Folder</td>
<td>controlaccessgeo</td>
<td>controlaccessgeo</td>
<td>required</td>
<td>CP = local</td>
<td>yes</td>
</tr>
</tbody>
</table>
Technical Specifications
Thank You

margaret.huang@philamuseum.org
Linked Open Data for a Comprehensive Numismatic Research Platform

Further Reading:

https://doi.org/10.5281/zenodo.1304270

https://goo.gl/MqMeGC

Ethan Gruber
Director of Data Science
American Numismatic Society
@ewg118

David Hill
Francis D. Campbell Librarian
American Numismatic Society
Silver Tetradrachm of Alexander III of Macedon, Amphipolis, 336 B.C. - 323 B.C. 1944.100.26704

Collection: MANTIS
http://numismatics.org/search

URI for this coin:
http://numismatics.org/collection/1944.100.26700
Research data:
Roman and Hellenistic Typologies linked to ~200,000 specimens in 30+ collections

Online Coins of the Roman Empire
http://numismatics.org/ocre
IGCH 1508

Findspot: 33.5625 (geo.lat) 35.36861111 (geo.long)
Closing date: 324 B.C. - 323 B.C.

1508 Salda (anc. Sidon), Phoenicia, 1829,
(881, 882, 884) 1852, 1863

Burial: c. 324/3 B.C. (Newell)*
Contents: 7200+ AV 3 separate pot hoards

Panticapaeum: 1 st.
Philiippa: 1 st. (de Luyres 1576)
Philip II: 302+ st.
Alexander III: 5600+ st. and some double
st.; many of Sidon (latest year 19) and
Ace (latest year 24)

Cius: 9 st. (BMC 1-2; de Luyres 2414-5;
Waddington 265; de Hirsch 1427;
McClellan 7458; Friedl.-Sallet 204;
Sotheby Feb. 19, 1868, 343)

Pergamus: 2 st. (Friedl.-Sallet 214; de
Luyres 2493)
Rhodes: 3 st. (Waddington 2770; Friedl.-
Sallet 234; BMC 10)

Pnytagoras: 2 st. (Waddington 4816)
Ptolemy I?: 1 st. (Svoronos pl. 1, 23)
Disposition: Istanbul 1800+ 430 (melted
down or not identifiable); Vienna 5
Phillips, 16 Alexanders; London 3;
Cambridge 1, Paris 8; Brussels 1;
Berlin 3; Sotheby June 1862 (Huber),
243-63; Dupré April 23, 1867, 317
(Pnytagoras); Sotheby Feb. 19, 1806,
343 (Cius); New York 17; Stockholm 2
Müller, Alexandre, p. 300, note 62
Waddington, RN 1865, pp. 3-25, Illus.; Melanges,

E.g., http://coinhoards.org/id/igch1508

The Inventory of Greek Coin Hoards

Coin hoard data: http://coinhoards.org

E.g., http://coinhoards.org/id/igch1508

Data Export

RDF/XML  KML  geoJSON
The American Numismatic Society: A Library, Archive, and Publisher
International Image Interoperability Framework (IIIF) – Standard APIs for image manipulation and display
Newell, Edward Theodore, 1886-1941

Toggle Names ▶

Jump to section: Relations Related Resources Annotations Associated Identifiers Export

Description

Exist Dates xEAC
1896 - 1941

Biographical or Historical Note

Edward Theodore Newell (1886-1941) was a U.S. numismatist. He served as the president of the American Numismatic Society between 1916 and 1941.

Edward T. Newell, the ANS's longest-serving President and perhaps the greatest numismatist of his generation, was born in Kenosha, Wisconsin in 1886. He attended Yale University from which he earned his bachelor of arts in 1907 and master of arts in 1909. It was during his tenure at Yale, that the young Newell — just 19 years old — first inquired in 1905 about membership in the ANS. By 1910, at the age of 24, Newell was elected to the Society's governing Council. In 1916 he was named ANS President, a position he would hold until his untimely death in 1941.

During his lifetime, Newell was considered to be the world's leading expert on the coinage of Alexander the Great and his successors. And Newell's collection — with more than 87,000 coins — was considered to be the largest private collection of Greek coins. Upon his death, Newell bequeathed this mammoth collection to the ANS. It was — and still is — the largest single donation given to the Society.

Chronology

- 1896: Born, Kenosha (Wis.)
- 1903 - 1905: Attended Yale, Yale University (Gree.)
- 1916: Received Archer M. Huntington Medal Award
- 1925: The Royal Numismatic Society awarded him its medal for "distinguished service in numismatic research."
- 1941: Died

Relations

Related Corporate, Personal, and Family Names

- org:memberOf
  - American Numismatic Society
- rel:colleagueOf
  - Miles, George Carpenter, 1894-1975
  - Newell, Adna
- x rebellion:With
  - Babelon, Jem., 1889-1978
  - Ball, George W.
  - Bronson, Campbell, 1876-1954
  - Clark, William E.
  - Dewey, William S.
  - Endicott, F. Munroe (Francis Munroe), 1879-1935
  - Ives, Herbert E. (Herbert Eugene), 1882-1953
  - Mears, Sawyer McArthur, b. 1905
  - Nieuw, Sydney P. (Sydney Philip), 1885-1969
  - Schultze, Hans P. A. T. (Hans Moritz Friedrich), 1913-1990
  - Wood, Howard, 1877-1938
  - Zerbe, Farren, 1871-1949

Published in xEAC

https://github.com/ewg118/xEAC

EAC-CPF
Authority records
Interoperable Models

EAD

MODS

TEI

RDF

Open Annotation
DC Terms, schema.org etc.

SPARQL Endpoint

RDF

Open Annotation
DC Terms, schema.org, etc.

Bio / W3C Org ontologies, SKOS, FOAF

EAC-CPF

Authorities
Enhancing research context: extracting related library/archival items via SPARQL
Extending the research ecosystem beyond the ANS: Pelagios and SNAC Integration

http://commons.pelagios.org/

http://snaccooperative.org/

ANS content
Challenges and Conflicts of Linked Data in Archives

Gregory Wiedeman, @GregWiedeman
University at Albany, SUNY
SAA 2018 Session 303
The Espy Project

- Collected documentation on American executions 1608-2002
- Index card summaries, 46 cu. ft. copies of original and published source material
- 1980s NSF grant to create “The Espy File” now in ICPSR
  - Main source for historical research on Capital Punishment in America
- Current CLIR grant to digitize and provide computational access to collection
Dear Mr. Epsy,

Mr. James did not live at Holly Road, the land at Cold Springs, north west of Cullman, please send this picture back.

With kind regards, I am

[Signature]

Mr. Matt Epsy, Jr.
P. O. Box 67
Headland, Alabama 36345

Dear Mr. Epsy,

The Probate Judge of Cullman County forwarded your letter of February 4 to me for reply.

I am a first cousin of George James, and complete information regarding his execution.

George was found guilty and was hanged on the complete story of the crime in my memory, of the actual hanging, newspaper articles and George a few minutes before his death. This last will & testament as well as a personal letter was given to Dr. Nagdory who was the at the hanging.

I will be happy to supply you with copies of writings to be "research". If your writ opposition to capital punishment, I would he for this purpose.

As a former peace officer, I feel we need

[Signature]

Route #2 Box 286
Hanceville, AL 35077
February 13, 1977

Mr. Matt Epsy, Jr.
P. O. Box 67
Headland, Alabama 36345

Dear Mr. Epsy,

the hotel was not able for any such insinuations, George was hanged on the theory, as we presume, that on account of inability to locate plaintiff’s suspicion was aroused, and they had the right to elect him for nonpayment of his bill. 22 C.C.J. 1975, Baile on Innkeepers and Hotels, e. 9. The case does not require a treatment of this question further than a statement of the general rule that when one has shown himself admitted as a guest the burden is on the innkeeper to justify ejectment of him, Baile, supra, § 101.

[1] If the question (objections to which were sustained) in assignments of error 10, and 11, should be canceled as not being material and relevant, we are of the opinion that no reversible error could be predicated upon this action of the court. The defendant made no offer to show by the house detective and page what was that the term, and the witnesses for defendant (including the clerks) were permitted to testify without objection, and without demeanor, that they had tried to find plaintiff, but were unable to do so; and the manager testified that he “went to his room a number of times, day and night,” and was unable to find him there, or anywhere else.

[2] JAMES v. STATE. (No. 59o.)

(Supreme Court of Alabama. June 30, 1915.

On Rehearing, July 2, 1915.)

1. Homograph (Mis)Reading in the First Degree—Evidence—Administrability.

ANDERSON, C. J., and McQUELAN and
SAYRE, JJ., concur.

JAMES v. STATE.
Metadata Matters

Occupation Field

- “Student”
- “Banana Dealer”
- “Beef Carrier”
- “Goat Herder”
- “Tiecutter”
- “Tourist”
- “Armed robber”
- “Asylum Escapee”
- “Bandit”
- “Criminal”
- “Cult Leader”
- “Gang Member”
- “Lunatic”
- “Male Nurse”
- “Retarded”
- “Slave”

- “Crime Convicted of” not “Crime Committed”
- Found that precision was often problematic in LOD vocabularies
- Create our own vocabulary?

VARIABLE DESCRIPTION LIST

V1  ICPSR Study Number-8451
V2  ICPSR Edition Number
V3  ICPSR Part Number-1
V4  Case Number
V5  Race of Offender
V6  Age at Execution
V7  Name of Offender
V8  Place of Execution
V9  Jurisdiction of Execution
V10 Crime Committed
V11 Method of Execution
V12 Date: Day
V13 Date: Month
V14 Date: Year
V15 Check Digit
V16 State of Execution
V17 County of Conviction
V18 ICPSR State Code
V19 Sex of Offender
V20 Compensation Case
V21 Occupation of Offender
Exposing Data in Context

Slave girl
(Need confirmation)
"A negro girl, convicted of the murder of her mistress, in New Orleans, La., was sentenced to be hung at 3 o'clock on the 26th ult. (6-26-1829). Subsequent to her trial, some new circumstances transpired, which induced a general belief that she was not guilty, or, if she was, that there was some accomplice more criminal than she, and the Governor was petitioned to reprieve the girl for three months, under the hope that she would be proved innocent, and the real murderer discovered. Accordingly on the forenoon of the day of execution, the Governor granted a reprieve for three months, which had scarcely been delivered to the sheriff, when the prisoner made a full confession of her crime to the jailer. As soon as this circumstance was made known to the Governor, he issued an order to the sheriff to carry the first sentence into effect at the
Ned, ditto, 29th Oct'r 1831, 425
Booker, ditto, 29th Oct'r 1831, 350
Shadrach, ditto, 29th Oct'r 1831, — 450
Squire, ditto, 29th Oct'r 1831, — 450
Frank, ditto, 29th Oct'r 1831, — 450 Killed in attempting to escape from jail.

Boson, ditto, 29th Oct'r 1831, — Total, 313, at $1
Solomon, ditto, 29th Oct'r 1831, 300
Jacob, Spottsylvania, 2d Nov'r 1831, 700
Davy, Southampton, 8th Nov'r 1831, 300
Jack Niles, Nansemond, 8th Nov'r 1831, 200
Dick, Westmoreland, 12th Nov'r 1831, 400
Frank, Southampton, 17th Nov'r 1831, — 600
Jim, ditto, 17th Nov'r 1831, — 300
Isaac, ditto, 17th Nov'r 1831, — 400
Stepney, Spottsylvania, 21st Nov'r 1831, — 150

Auditor's 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831.

A valuable document but not perfect. See notations regarding errors.


Both executed 9/12/31. See master list

Jack Niles (Slave) Executed, Nansemond Co., Va. Owners reimbursed $900 on 11-8-1831

Executed 9/23/31. See master list

Davy
Executed (Land); Southampton Co., Va. Owners reimbursed $300 on 11-8-1831

Nashville, Nat Turner's slave insurrection

Executed 9/7/31 see master list.
Does Linked Data align with the Archival Mission?

- Archives use context to efficiently provide access to unique material at scale
- Researchers will not use SPARQL to access this data
- In practice, do URIs provide or obscure context?
- Does a Linked Data URI convey false objectivity or false authority?
- Does this scale?

- As we start seeing archival collections themselves as data, Linked Data may not be a good fit
Challenges and Conflicts of Linked Data in Archives

Gregory Wiedeman, @GregWiedeman
University at Albany, SUNY
SAA 2018 Session 303
Thank you!

- Megan Burdi, Archivist for Digital Initiatives, Archives of American Art, Smithsonian Institution, burdim@si.edu
- Katherine Barbera, Assistant Archivist, Carnegie Mellon University, kbarbera@andrew.cmu.edu
- Matthew R. Miguez, Metadata Librarian, Florida State University, mmiguez@gsu.edu
- Samantha Norling, Digital Collections Manager, Newfields, snorling@discovernewfields.org
- Seth Shaw, Application Developer, University of Nevada, Las Vegas, University Libraries, seth.shaw@unlv.edu
- Mark Matienzo, Collaboration & Interoperability Architect, Stanford University, matienzo@standford.edu
- Toby Reiter, Web Developer, Archives of American Art, Smithsonian Institution, reitert@si.edu
- Elizabeth Ehrnst, Digital Initiatives Librarian, Georgia O'Keeffe Museum, eehrnst@okeeffemuseum.org
- Margaret Huang, Digital Archivist, Philadelphia Museum of Art, Margaret.Huang@philamuseum.org
- Ethan Gruber, Director of Data Science, American Numismatic Society, gruber@numismatics.org
- Greg Wiedeman, University Archivist, University at Albany, SUNY, gwiedeman@albany.edu

Archives and Linked Data Interest Group: https://archival.github.io/