Sharing LOUDly

Linked Open Usable Data at the J. Paul Getty Trust

Lily Pregill | Systems Architect | lpregill@getty.edu

8th International Conference of Art Libraries, Amsterdam, Oct 4-5, 2018
Getty Digital will through increased engagement and collaboration, enable the programs and broader cultural heritage community to create, manage and publish high quality, interoperable and easy-to-use digital content in a coherent and sustainable platform.

Our major areas of concentration are digitization, linked open data, enabling computational art history research, and building a shared, distributed data environment.
WHY LOD?

Data Management
- More accurate and explicit descriptions
- Easier system integration

Data Publication
- Share what we know
- Let others build on that knowledge

Data Consumption/Exchange
- To more easily make use of our own data
- You have valuable data, too!
WHY LOD?
LOD VISION

LOD ECOSYSTEM

CONSISTENT ACCESS

GETTY CONSERVATION INSTITUTE + GETTY FOUNDATION + GETTY RESEARCH INSTITUTE + J. PAUL GETTY MUSEUM
OVERVIEW

1. Identities > Getty Vocabularies
2. Data model > Linked Art
3. System > Arches
GETTY VOCABULARIES

- AAT, ULAN, and TGN released as LOD 2014
- Now beta testing two new LOD services:
  1. AAT JSON-LD web service
  2. OpenRefine reconciliation service
AAT JSON-LD WEB SERVICE
OPENREFINE RECONCILIATION SERVICE

transform strings to things

“Peter Paul Rubens”

<http://vocab.getty.edu/ulan/500002921>

Pietro Paulo Rubens

P. P. Rubbens

بيتر بول روبنس

Petrus Paulus Rubens

ピーテル・パウル・ルーベンス
LINKED ART → https://linked.art

Linked Art is a community working together to create a shared data model based on Linked Open Data to describe art.

- Developed by Rob Sanderson, Semantic Architect, and David Newbury, Software Architect at the J. Paul Getty Trust

- Grew out of work with the American Art Collaborative (AAC)
WHY LINKED ART?
USABLE vs COMPLETE
WHY LINKED ART? USABLE TO HUMANS AND MACHINES

```
{
  "@context": "https://linked.art/ns/v1/linked-art.json",
  "id": "https://data.getty.edu/museum/11733",
  "type": "ManMadeObject",
  "label": "Attic Black-Figure Neck Amphora",
  "classified_as":
   [{"id": "aat:300148696", "type": "Type", "label": "Amphora"}]
  "made_of":
   [{"id": "aat:300015045", "type": "Material", "label": "Terracotta"}]
  "produced_by":
   [{"id": ".../create", "type": "Production", "label": "Creation event"}
    // ...
}
```
LINKED ART DESIGN PRINCIPLES

- **Usability**, not 100% precision / completeness
- Consistently solves actual challenges from real data
- Development is iterative, as new use cases are found
- Solve 90% of use cases, with 10% of the effort
LINKED.ART MODEL FUNDAMENTALS

- Profile of CIDOC-CRM
  Create recipes for modeling art-related events

- Getty Vocabularies are used as core identities
  Leverage 30 years of thesauri and authority work

- JSON-LD primary target serialization
  Meet developers where they work
LINKED ART GOALS

- Robust community participation and adoption
- Usable data by machines and users (developers)
- Data interoperability across the domain
- Standard model for vendor data management systems to adopt
- A LOD future attainable by all
LINKED ART COLLABORATION

- Getty
- Rijksmuseum
- Louvre
- Metropolitan Museum of Art
- Smithsonian
- MoMA
- National Gallery of Art
- Philadelphia Art Museum
- Indianapolis Art Museum
- Harvard University
- Princeton University
- Yale Centre for British Art
- Oxford University
- Academica Sinica
- ETH Zurich
- FORTH
- Zeri Foundation (U. Bologna)
- Canadian Heritage Info. Network
- American Numismatics Society
- Frick Art Reference Library
WHY LINKED ART?

“Simple things should be simple; complex things should be possible.”

Alan Kay
ARCHES → https://www.archesproject.org/

- Open source data management platform for cultural heritage; originally for immovable heritage

- Project of Getty Conservation Institute and World Monuments Fund; released 2013

- CRM integration allows for mapping of data to classes and properties using a graph structure
ARCHES NEW USES

- Provenance Index
- Ed Ruscha Streets of Los Angeles
- Art and Archaeology Technical Abstracts (AATA)
- DISCO: Data Integration for Conservation Science
- PhotoTech
RESOURCES

- Getty Vocabularies: http://www.getty.edu/research/tools/vocabularies/
- Linked Art: https://linked.art
- Arches: https://www.archesproject.org/
- Provenance Index Rebuild:
  http://www.getty.edu/research/tools/provenance/provenance_remodel/index.html
  http://blogs.getty.edu/iris/tags/getty-provenance-index/
- Ruscha: http://www.getty.edu/research/scholars/digital_art_history/ruscha/index.html
- AATA: http://blogs.getty.edu/iris/the-database-conservators-built/
- DISCO: http://www.getty.edu/conservation/our_projects/science/integrating_data/
ACKNOWLEDGEMENTS

Thanks to following folks for wisdom, slides, and icons:

Rob Sanderson, Semantic Architect, J. Paul Getty Trust
David Newbury, Software Architect, J. Paul Getty Trust
Melissa Gill, Metadata Lead, Getty Research Institute

Compass by Setyo Ari Wibowo from the Noun Project
Share by Vectorstall from the Noun Project
THANK YOU - BEDANKT!

lpregill@getty.edu | @technelily

GETTY CONSERVATION INSTITUTE + GETTY FOUNDATION + GETTY RESEARCH INSTITUTE + J. PAUL GETTY MUSEUM