Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Linked.Art & Vocabularies: Linked Open Usable Data

Rob Sanderson, David Newbury
Semantic Architect, Software & Data Architect
J. Paul Getty Trust

rsanderson@getty.edu, dnewbury@getty.edu
@azaroth42, @workergnome

Linked Art & Vocabularies

RDF & Linked Data & Ontologies & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

What is RDF?

"RDF is a standard model for data interchange on the Web."

- A way for computers to work with facts
- A way to express statements about resources
- A W3C (web-based, web-friendly) data standard

Subject -> Predicate -> Object

Linked Art & Vocabularies

What is Linked Data?

"The term Linked Data refers to a set of best practices for publishing structured data on the Web."

- Some rules that help make RDF useful for others
- Every ID is a HTTP URL, and it returns info about itself
- You should use other people's URLs whenever you can

Subject (URI) -> Predicate (URI) -> Object (URI)

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

What are Ontologies?

"The formal naming and definition of the types, properties and interrelationships of the entities that exist for a particular domain."

- The application of the LOD framework to a domain
- Ontologies provide meaning to data
- Ontologies are most useful when used by multiple organizations and datasets

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

Cultural Heritage Ontologies?

"The CIDOC-CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation."

- An event-based way to look at the world
- Mature ontology in Museum domain
- "Semantic glue" between different information sources
- Highly theoretical

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

What are Vocabularies?

"A controlled vocabulary is an organized arrangement of words and phrases used to index content and/or retrieve content through browsing or searching."

- Shared terminology to reduce ambiguity
- In LOD, vocabularies provide reusable identities for shared entities across datasets
- Vocabularies are hubs in the web of data, allowing easy interconnections between systems

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Linked Art

https://linked.art/

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art Model

A Linked Open Data model, designed to work across many museums and to enable functional applications.

- Homogenized model across many institutions
- Supports varying levels of completeness
- Enables “round-tripping” of data
- Framework: Linked Open Data
- Ontology: CIDOC-CRM (plus a little)
- Vocabularies: AAT, ULAN, TGN
- Focused on Usability

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art Model

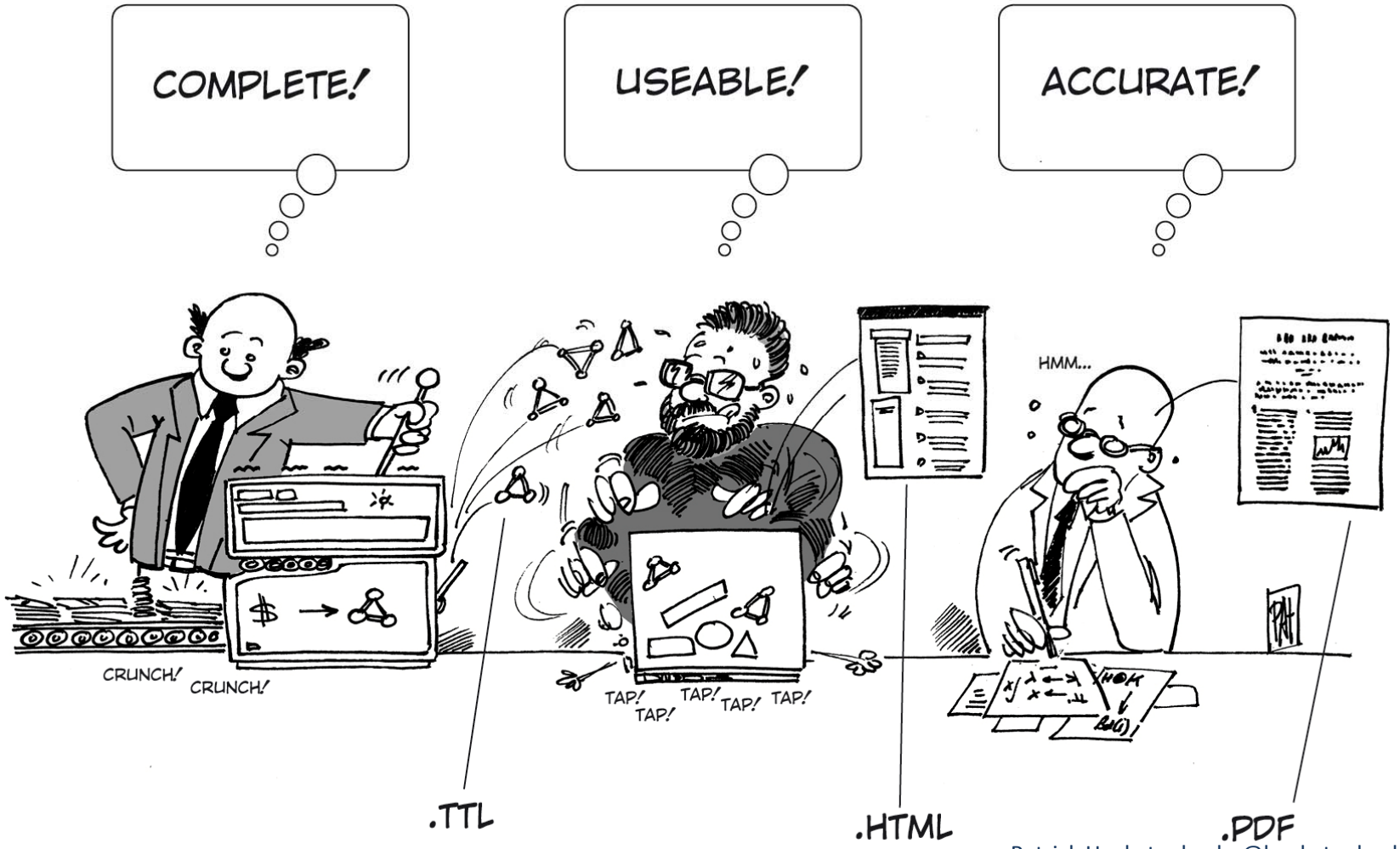
Linked Art provides patterns and models that enable cultural heritage institutions to easily publish their data for use by both event-based digital research projects and for non-specialist developers.

90% of the use cases of
90% of the organizations, with
10% of the complexity

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu



Patrick Hochstenbach, @hochstenbach

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

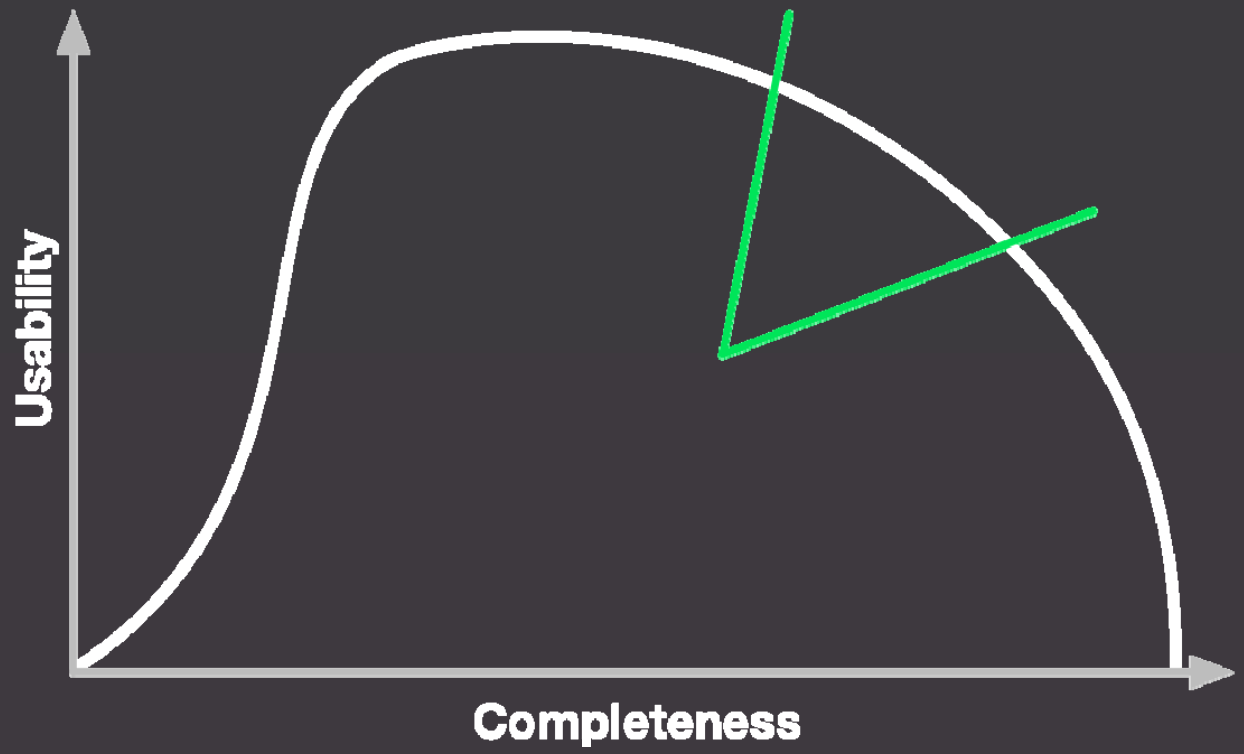
Usable vs Complete

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Target Zone for Scope

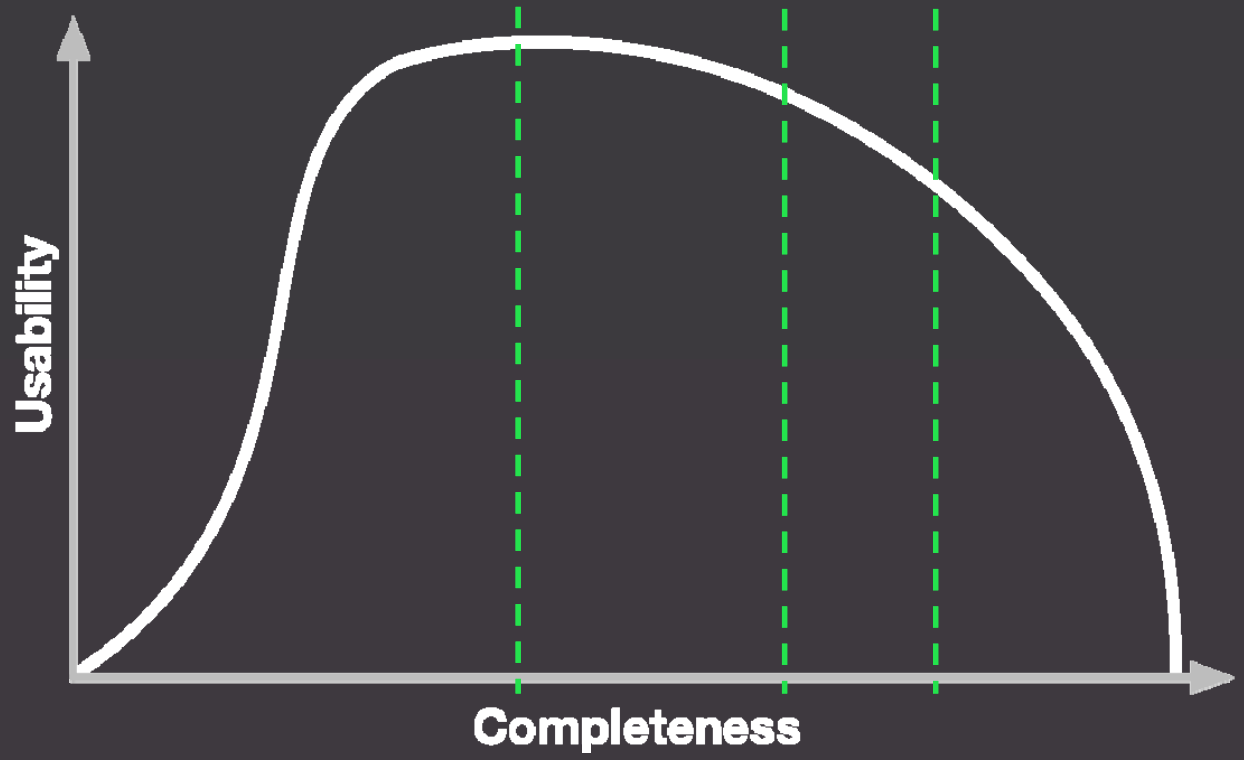


Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Incremental Complexity



Linked Art & Vocabularies

Successful Data Models…

- Solve actual challenges, documented as use cases
  - Using data that is captured and available
- Allow consistent description of shared use cases
  - Allow for addition of further information
- Can be productively used
  - Via easy-to-implement services
  - With easy-to-implement applications
- Provide interoperability with other data
- Are clearly documented with relevant examples

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

Successful Data Models…

Are developed …

- Iteratively
  - We will not get it right first time
- Responsively
  - We will adapt it in response to feedback
- Responsibly
  - We will consider changes/features carefully
- Collaboratively
  - We will engage with the community and stakeholders

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

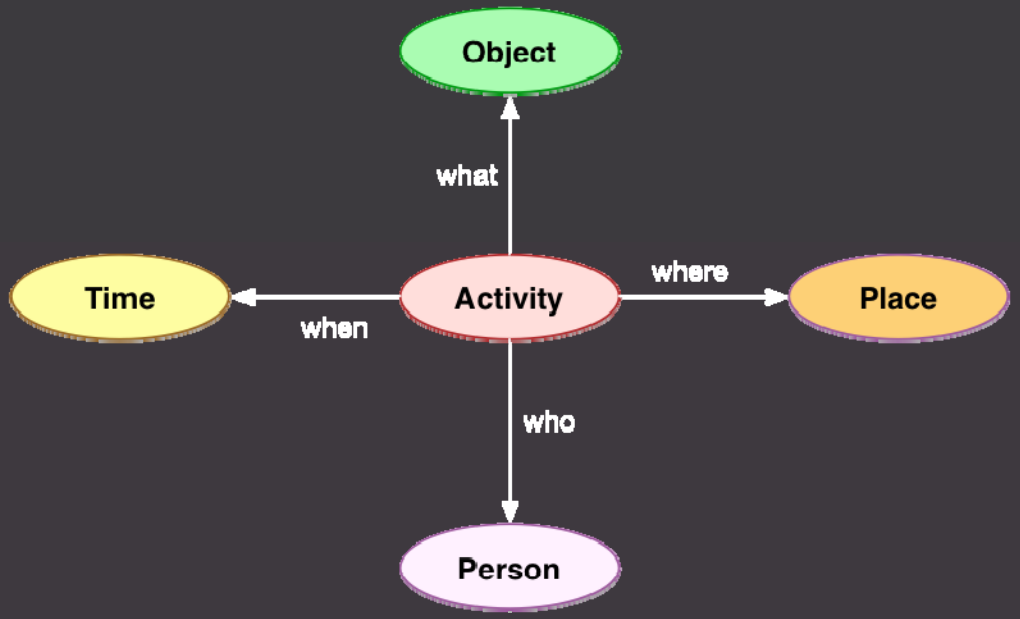
Linked Art & Vocabularies

Successful Data Models…

Focus on the primary resources of concern for the domain

Core Classes:

- Object
- Activity
- Person
- Place
- Time



@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Successful Data Models…

Are Consistent and Unsurprising

The more consistency across the model, the less you need to learn and remember.

Common Patterns designed and used:

- Consistent naming scheme
- Classification of Specifics
- Partitioning
- Statements about Resources

Linked Art & Vocabularies

Pattern: Classification of Specifics

We use AAT to provide more specific classifications of resources than are possible using CIDOC-CRM

- Object: Painting vs Sculpture
- Identifier: Accession Number vs Database Id
- Place: Country vs City
- Group: Museum vs Art Dealer
- Activity: Engraving vs Casting

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Pattern: Partitioning

Many types of resource can be partitioned, creating identity for further description of more granular parts

- Objects: Parts, Features with different dimensions etc
- Activities: Exhibition vs Venue, Auction vs Lot
- Locations: Country, Region, City, District, Building
- Texts: Set, Volume, Chapter, Page
- Organizations: Institution, Program, Department

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Pattern: Statements about Resources

Resources can have statements with classifications from AAT for the topic, co-existing with machine readable data.

- Materials
- Provenance
- Attribution
- Biography
- Description

Linked Art & Vocabularies

Focus: People and Places

AAT provides a rich list of identities for classification.
ULAN and TGN provide identities for actors and locations.

Fill different needs within the model:

- Classification important for shared semantics
- Person/Place important for shared entities

Provides a hub in the web of data for connecting across organizations, without prior agreement

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

Vocabularies and CIDOC-CRM?

New Challenge: Vocabularies use different ontologies!

AAT: Minimal impact - labels
ULAN, TGN: Significant impact - entire entity model

How do we ensure consistency and usability of the Vocabularies in the Linked Open Data ecosystem?

@azaroth42
@workergnome
rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

APIs & Profiles & Formats & Applications

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

API: Agreement Preceding Interaction*

The same data can be exposed in different ways to different audiences simultaneously.

- Need to understand how to access the data
- APIs can provide different functionality
- Fewer ways with more users is more sustainable
- Web provides HTTP protocol for data transport

(* API is really: Application Programming Interface)

Linked Art & Vocabularies

Data Profiles

Need the shape of the data to be consistent to be easy to use. Different audiences thus need different shapes.

- Current profile (SKOS-XL, custom GVP) for Researchers
- Add Linked Art profile for Museum audience
- Add schema.org profile for Web/SEO audience

Different representations or descriptions of the same entity, not different entities.

@azaroth42
@workergnome
rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

Developer-Friendly Format

```
{
  "@context": "https://lod.museum/ns/context/1/full.jsonld",
  "id": "https://lod.museum/example/object/1",
  "type": "ManMadeObject",
  "classified_as": "aat:300033618", # by reference
  "label": "Example Painting",
  "made_of": {
    "id": "aat:300015045", # by (minimal) value
    "type": "Material",
    "label": "watercolor"
  }
}
```

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Applications

Developers that can rely on consistent format, shape and interactions, can build robust and innovative applications.

- Need many such applications to validate usability
- Notifications of updates to remain synchronized
- Notifications of contributions to dataset
  - … via reference to external entities?

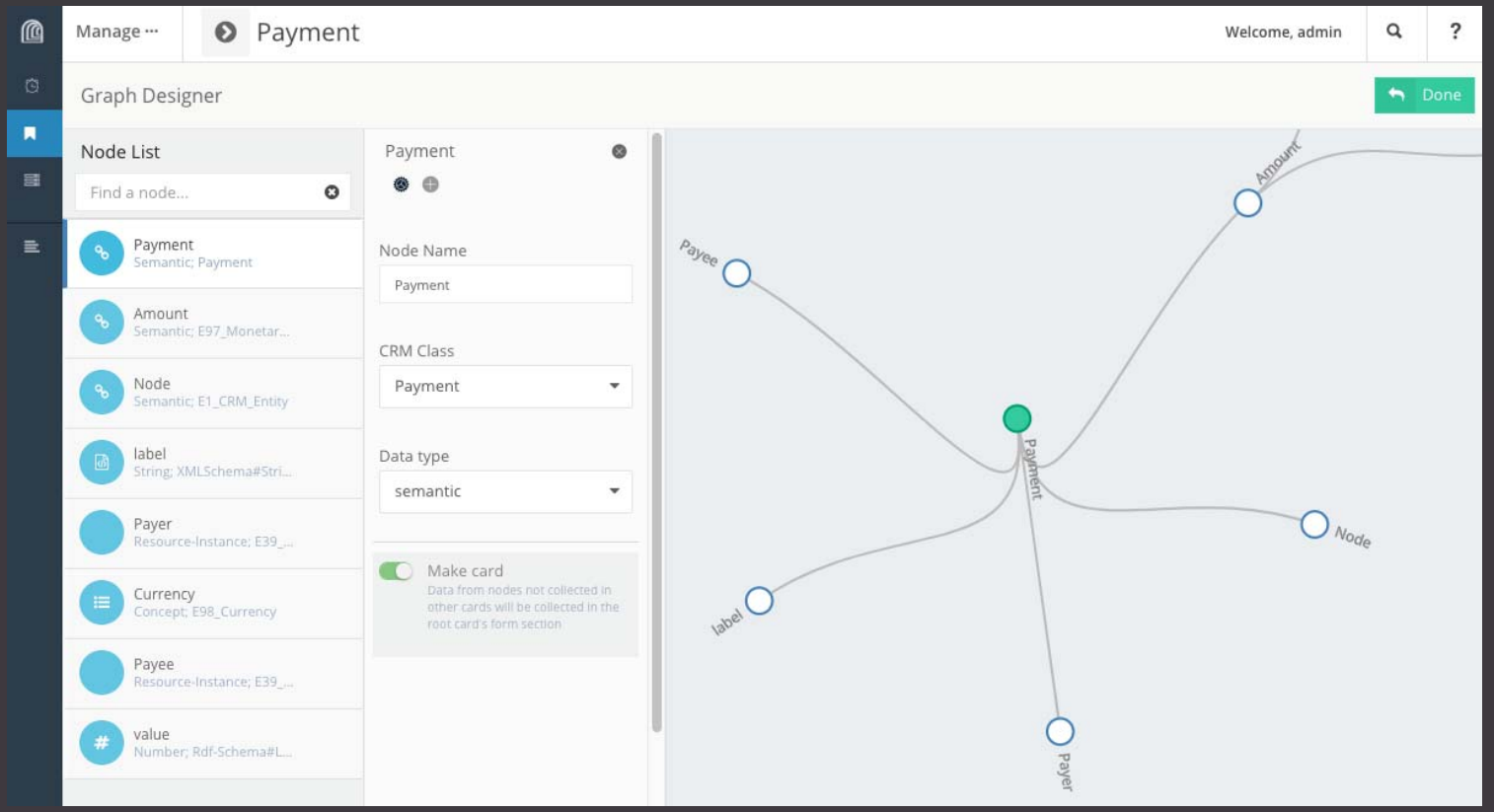
Translations provide easy internationalization for applications - a huge benefit to the community

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Arches

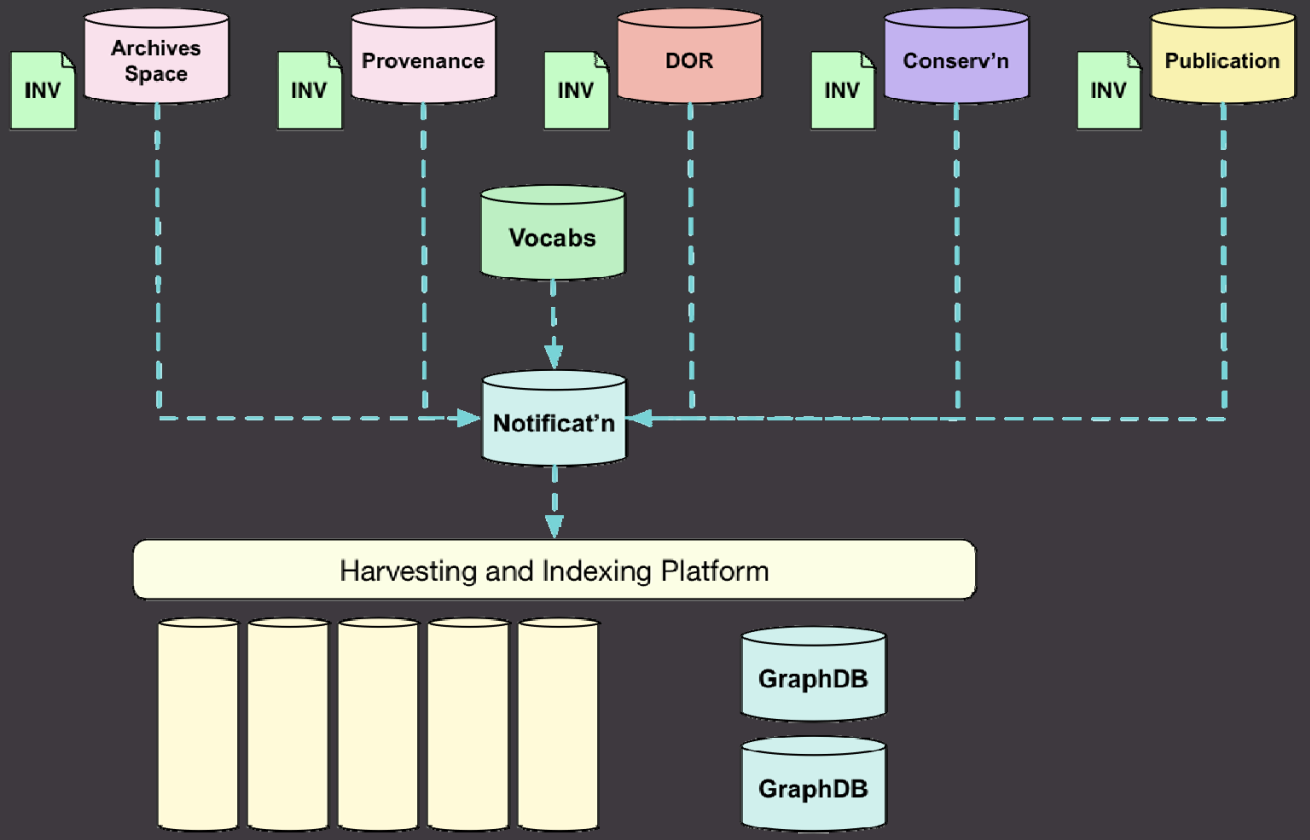


Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Notification Infrastructure



Linked Art & Vocabularies

Summary

- Linked Open Data!
- Vocabularies are Identity Providers
- Linked Art – usability focused data model
  - Uses CIDOC-CRM and Vocabularies
- Technical improvements suggested for improving consistency and functionality

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

Thank You!

Rob Sanderson, David Newbury
rsanderson@getty.edu, dnewbury@getty.edu
@azaroth42, @workergnome

@azaroth42
@workergnome

rsanderson; dnewbury @getty.edu

Linked Art & Vocabularies

@azaroth42
@workergnome

rsanderson;
dnewbury
@getty.edu

Discuss!