

Getty Vocabularies Technical Update CIDOC 2020

Gregg Garcia
Software Architect, Getty Digital

Getty Vocabularies Publications and Services

Full XML and Tab-delimited Relational Tables Downloads

- <http://www.getty.edu/research/tools/vocabularies/obtain/download.html>

Vocabulary Web Site

- <http://www.getty.edu/research/tools/vocabularies/>
- Human-readable search, hierarchy and full-record display

XML Web Services

- http://www.getty.edu/research/tools/vocabularies/vocab_web_services.pdf

Reconciliation Service for OpenRefine

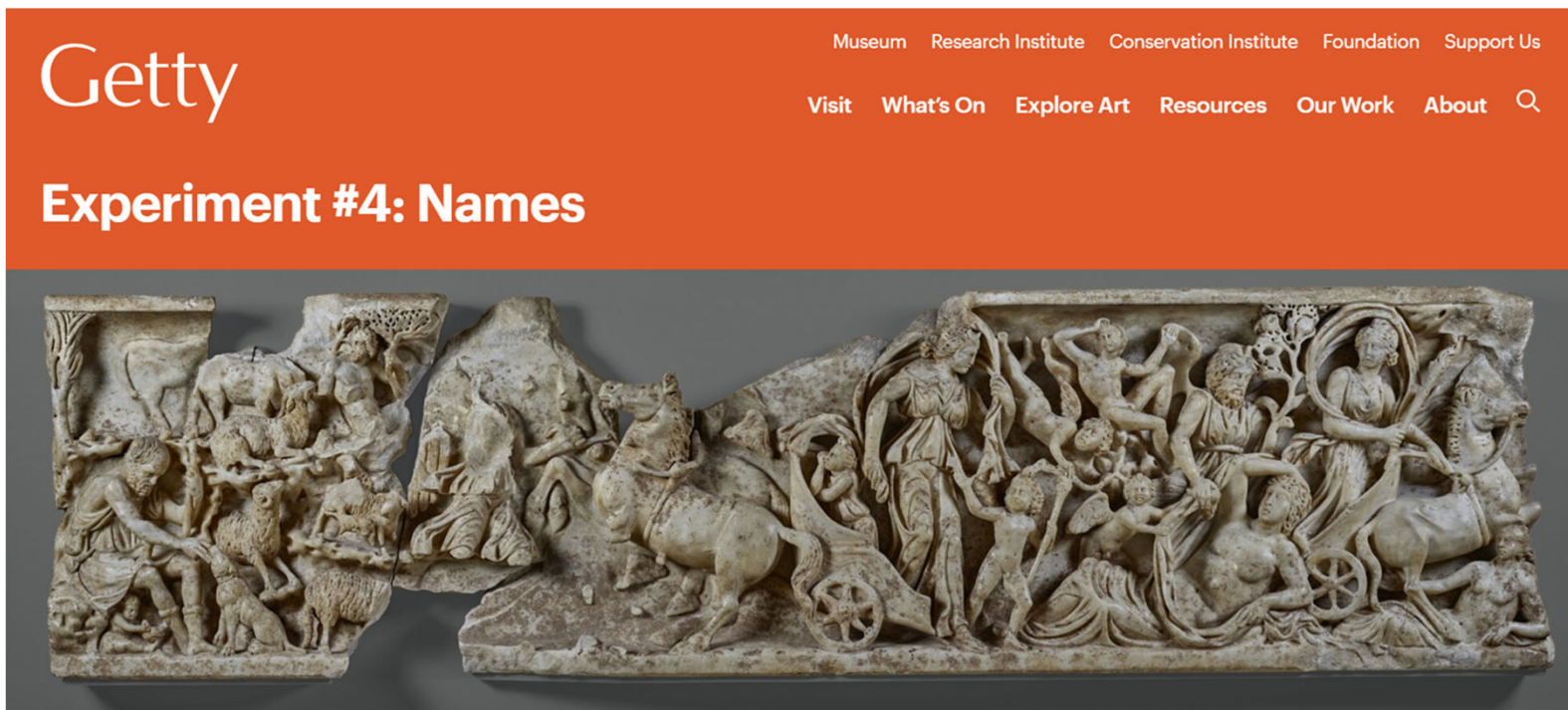
- <http://services.getty.edu/vocab/reconcile/>
- Accessible using OpenRefine tool or HTTP

Linked Data Site with Downloads and SPARQL Endpoint

- <http://vocab.getty.edu>

Use Case: Getty Names Experiment

<https://experiments.getty.edu/names/>



Experiment Objectives


- Create an application in one day
- Find relationships between people in Getty data sets
- Use ULAN as the primary authority
- Return “well-known” people in search
 - Primary criteria is a text search
 - Algorithm boosts results for people who are cited the most in our data
- Use Wikidata coreferences for images

Step 1: Query Data for Search Index

```
SELECT * WHERE {  
  BIND(<URI> as ?s)  
  ?s gvp:prefLabelGVP/skosxl:literalForm ?prefName.  
  OPTIONAL { ?s foaf:focus/gvp:biographyPreferred/schema:description ?bio. }  
  ?s rdfs:label ?labels.   
  OPTIONAL { ?s dcterms:source ?sources. ?sources a bibo:Document}   
  OPTIONAL { ?s skos:exactMatch ?corefs FILTER NOT EXISTS {  
    FILTER(CONTAINS(str(?corefs),'wikidata')) } }  
}
```






GVP Preferred Name →
GVP Preferred Bio ↓
All Name Variants ←
All Sources ←
All Coreferences Excluding Wikidata ←



Step 2: Get Wikidata Links

 Wikidata Query Service

[Examples](#) [Help](#) [More tools](#)

```
1 select * {  
2   ?wd wdt:P245 ?ulan  
3 }  
4
```

  **91712 results in 2427 ms** [Code](#) [Download](#)

Step 3: Load Final JSON into Search Index

```
{
  "uri" : "http://vocab.getty.edu/ulan/500115588",
  "display_text" : "Gogh, Vincent van (Dutch painter and draftsman, 1853-1890)",
  "names" : [
    "Vincent van Gogh",
    "梵高",
    "ビンセントゴッホ",
    "וינסנט ואן גוך",
    "Gogh, Vincent Willem van",
    "Fan-ku",
    "Gogh, Vincent Van",
    "ゴッホ",
    "וינסנט ואן גוך",
    "Fangu",
    "Fan-kao",
    "van Gogh, Vincent",
    "Gogh",
    "Van Gogh, Vincent",
    "j. van gogh",
    "Gogh, Vincent van",
    "Gogh, Vincent-Willem van",
    "Fan'gao",
    "van gogh",
    "Fangu, Wensheng",
  ],
  "coreferences" : {
    "count" : 2,
    "values" : [
      "https://www.wikidata.org/wiki/Q5582",
      "http://id.loc.gov/authorities/names/n79022935"
    ]
  },
  "citations" : {
    "count" : 12,
    "values" : [
      "http://vocab.getty.edu/ulan/source/2100182948",
      "http://vocab.getty.edu/ulan/source/2100133206",
      "http://vocab.getty.edu/ulan/source/2100039894",
      "http://vocab.getty.edu/ulan/source/2100181343",
      "http://vocab.getty.edu/ulan/source/2100043316",
      "http://vocab.getty.edu/ulan/source/2100182949",
      "http://vocab.getty.edu/ulan/source/2100041444",
      "http://vocab.getty.edu/ulan/source/2100045787",
      "http://vocab.getty.edu/ulan/source/2100154640",
      "http://vocab.getty.edu/ulan/source/2100181667",
      "http://vocab.getty.edu/ulan/source/2100042266",
      "http://vocab.getty.edu/ulan/source/2100041042"
    ]
  },
  "wikilink" : "https://www.wikidata.org/wiki/Q5582"
}
```

Search Results for “gogh”

Results



Gogh, Vincent van (Dutch painter and draftsman, 1853-1890)

Also known as Vincent van Gogh, 梵高,
ビンセントゴッホ, גוג, וינסנט ואן, Gogh,
Vincent Willem van, Fan-ku, Gogh,
Vincent Van, ゴッホ, וינסנט ואן, Fangu,
Fan-kao, van Gogh, Vincent, Gogh, Van
Gogh, Vincent, j. van gogh, Gogh,
Vincent van, Gogh, Vincent-Willem van,
Fan'gao, van gogh, Fangu, Wensheng,
Van-Gog, Vintsent, v. van gogh, Vincent
Van Gogh



Gogh-Bonger, Johanna van (Dutch collector, 1862-1925)

Also known as Gogh-Bonger, Johanna
van, Bonger, Jo, Van Gogh-Bonger,
Johanna, Gogh-Bonger, Johanna von,
Johanna van Gogh-Bonger, Johanna
Gesina Gogh-Bonger, Gogh-Bonger, J.
van, Gogh-Bonger, Johanna Gesina,
Gogh, Jo van, Gogh-Bonger, Johanna
Gesina van, Bonger, Johanna van Gogh-



Du Quesne-van Gogh, Elisabeth Huberta (Dutch artist, 1859-1936)

Also known as Gogh, Elisabeth Huberta
van, Du Quesne-van Gogh, Elisabeth
Huberta, Elisabeth Huberta du Quesne-
van Gogh



Van Gogh Museum (Dutch museum, Amsterdam, contemporary)

Also known as Van Gogh Museum,
Amsterdam, Rijksmuseum Vincent van
Gogh, Amsterdam, Rijksmuseum Vincent
van Gogh, Van Gogh Museum,
Rijksmuseum Vincent van Gogh
Amsterdam, Stichting Van Gogh
Museum

Step 4: Visualization of ULAN Associative Relationships

Getty

Museum Research Institute Conservation Institute Foundation Support Us

Visit What's On Explore Art Resources Our Work About 

Some of this person's relationships:



Student of
Cormon, Fernand (French painter and teacher, 1845-1924)
[Person URI](#)
[WikiData](#)

Getty

Possible Next Steps

- Add people and relationships from other Getty data sets
- Add links to Getty resources
- Add links to more external authorities (LoC, VIAF, etc.)
- Add More through the magic of linked data

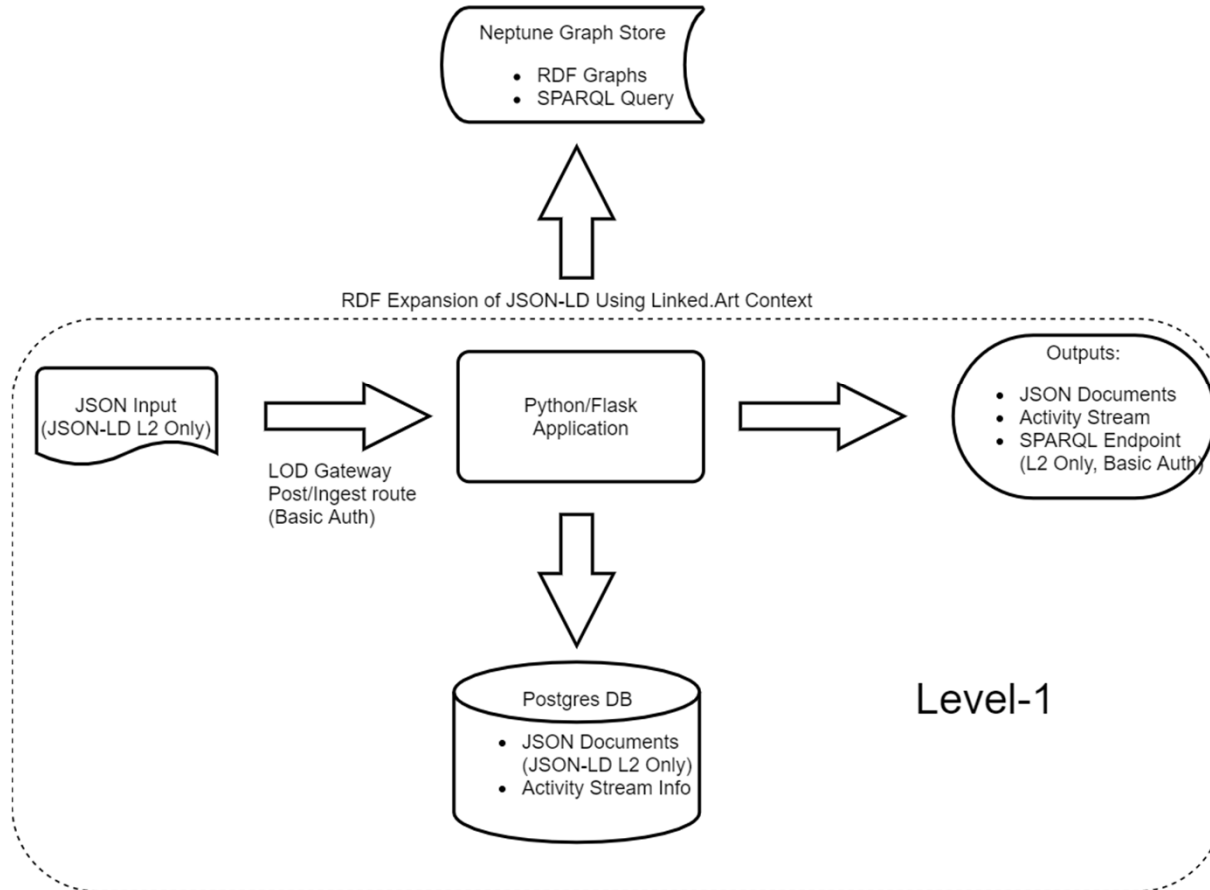
The Getty Vocabularies as Linked.Art Update

Why Linked.Art?

- Institutional effort to bring all data sets at the Getty under a unified data model
- Linked.Art provides a bridge for representing the Vocabularies in the CIDOC CRM
- JSON-LD offers a simple serialization format for developers with the ability to expand into linked data triples
- The GDi LOD Gateway provides a platform for storing JSON-LD documents that are expanded into full RDF on load

GDi LOD Gateway

version 1.0, Gregg Garcia, 8/18/20



Current State of the Getty Vocabularies LOD

- AAT, ULAN and TGN subjects represented as SKOS concepts
- ULAN and TGN biographical and place data represented in Schema.org through foaf:focus property
- Revision history represented by PROV ontology
- Lucene index search on terms, identifiers and scope notes
- RDF inferencing done by GraphDB backend

New Linked.Art Version of Getty Vocabularies LOD

- AAT entity represented as CRM Type
- ULAN entity represented as CRM Person or Group
- TGN entity represented as CRM Place
- Compact JSON-LD format provides aliased classes and properties for simple, in-code processing
- Activity stream provides a history of created, updated, and deleted records
- Neptune graph store backend for SPARQL

SPARQL Query Comparison - All Terms Example for “Rhyta”

SKOS:

```
SELECT ?labels WHERE {  
  <http://vocab.getty.edu/aat/300198841> rdfs:label ?label .  
  BIND(str(?label) as ?labels)  
}
```

CRM:

```
SELECT ?labels WHERE {  
  { <https://data.getty.edu/vocab/aat/300198841> crm:P1_is_identified_by ?nm_uri.  
    ?nm_uri a crm:E33_E41_Linguistic_Appellation.  
    ?nm_uri crm:P190_has_symbolic_content ?labels. }  
}
```


Entity JSON Comparison - Single Term Example for “Rhyta”

Current:

```
{
  "Subject" : {
    "type" : "uri",
    "value" : "http://vocab.getty.edu/aat/300198841"
  },
  "Predicate" : {
    "type" : "uri",
    "value" : "http://www.w3.org/2000/01/rdf-schema#label"
  },
  "Object" : {
    "xml:lang" : "en",
    "type" : "literal",
    "value" : "rhyta"
  }
}
```

Linked.Art:

```
{
  "Identified_by" : [
    {
      "id": "http://vocab.getty.edu/aat/term/1000198841-en",
      "type": "Name",
      "content": "rhyta",
      "language": [
        {
          "id": "http://vocab.getty.edu/language/en",
          "type": "Language",
          "_label": "en"
        }
      ]
    }
  ]
}
```

Entity JSON Comparison - Single Term Example for “Rhyta”

Current:

```
{
  "Subject" : {
    "type" : "uri",
    "value" : "http://vocab.getty.edu/aat/300198841"
  },
  "Predicate" : {
    "type" : "uri",
    "value" : "http://www.w3.org/2004/02/skos/core#prefLabel"
  },
  "Object" : {
    "xml:lang" : "en",
    "type" : "literal",
    "value" : "rhyta"
  }
}
```

Linked.Art:

```
{
  "Identified_by" : [
    {
      "id": "http://vocab.getty.edu/aat/term/1000198841-en",
      "type": "Name",
      "content": "rhyta",
      "language": [
        {
          "id": "http://vocab.getty.edu/language/en",
          "type": "Language",
          "_label": "en"
        }
      ]
    }
  ]
}
```

Entity Model Comparison - Single Term Example for “Rhyta”

Vocab XML:

```
<Terms>
  <Preferred_Term>
    <Term_Text>rhyta</Term_Text>
    <Term_ID>1000198841</Term_ID>
    <Term_Languages>
      <Term_Language>
        <Language>70051/English</Language>
      </Term_Language>
    </Preferred_Term>
</Terms>
```

Linked.Art:

```
{
  "Identified_by" : [
    {
      "id": "http://vocab.getty.edu/aat/term/1000198841-en",
      "type": "Name",
      "content": "rhyta",
      "language": [
        {
          "id": "http://vocab.getty.edu/language/en",
          "type": "Language",
          "_label": "en"
        }
      ]
    }
  ]
}
```

SPARQL Query for ULAN Search Index (Current LOD)

```
SELECT * WHERE {  
  BIND(<URI> as ?s)  
  ?s gvp:prefLabelGVP/skosxl:literalForm ?prefName.  
  OPTIONAL { ?s foaf:focus/gvp:biographyPreferred/schema:description ?bio. }  
  ?s rdfs:label ?labels.   
  OPTIONAL { ?s dcterms:source ?sources. ?sources a bibo:Document}   
  OPTIONAL { ?s skos:exactMatch ?corefs FILTER NOT EXISTS {  
    FILTER(CONTAINS(str(?corefs),'wikidata')) } }  
}
```

GVP Preferred Name →

GVP Preferred Bio ↓

← **All Name Variants**

← **All Sources**

← **All Coreferences Excluding Wikidata**

SPARQL Query for ULAN Search Index (Linked.Art/CRM)

```
SELECT * WHERE {  
  BIND(<URI> as ?s)  
  GRAPH ?s {  
    ?s rdfs:label ?prefName.  
    ?s crm:P67i_is_referred_to_by/crm:P190_has_symbolic_content ?bio.  
    ?s crm:P1_is_identified_by/crm:P190_has_symbolic_content ?labels.  
    OPTIONAL { ?s crm:P1_is_identified_by/crm:P67i_is_referred_to_by ?src_xref.  
      ?src_xref crm:P1i_identifies ?sources.  
      ?src_xref crm:P2_has_type <http://vocab.getty.edu/aat/300311936> }  
    OPTIONAL { ?s skos:exactMatch ?corefs FILTER NOT EXISTS { FILTER(CONTAINS(str(?corefs),'wikidata')) }  
  }  
}
```

GVP Preferred Name

GVP Preferred Bio

All Name Variants

All Sources

All Coreferences Excluding Wikidata

Use of the Getty Vocabularies (AAT) for Terminology within all Getty Profiles

From <https://linked.art/model/>:

- The Getty Vocabularies as core sources of identity for domain-specific terminology
 - Core terminology is part of the application profile documentation directly, when interoperability between systems relies on the selection.

Using our own authority to describe our data

Getty Profiles Title Example

Museum Collection:

```
{
  "id": "https://data.getty.edu/museum/collection/object/c88b3df0-de91-4f5b-a9ef-7b2b9a6d8abb/name/67c4be35-3098-47e2-bdcc-02556a62614f",
  "type": "Name",
  "_label": "Primary Title",
  "content": "Irises",
  "classified_as": [
    {
      "id": "http://vocab.getty.edu/aat/300417193",
      "type": "Type",
      "_label": "Titles (General, Names)"
    },
    {
      "id": "http://vocab.getty.edu/aat/300404670",
      "type": "Type",
      "_label": "Preferred Term"
    }
  ]
},
```

Getty Profiles Title Example

Research Collections Viewer:

```
{
  "id": "https://data.getty.edu/research/collections/component/65940213-1c81-5201-b582-bac076496444/Title",
  "type": "Name",
  "classified_as": [
    {
      "id": "http://vocab.getty.edu/aat/300417201",
      "type": "Type",
      "_label": "Title"
    },
    {
      "id": "http://vocab.getty.edu/aat/300404670",
      "type": "Type",
      "_label": "preferred terms"
    }
  ],
  "content": "Pacific Coast Highway"
},
}
```


Next Steps

- API documentation
- Model refinement based on internal Getty use-cases
- New, streamlined contribution methods
- Better reconciliation algorithms
- More coreferences

Thank You, CIDOC!