# GVP LOD: ONTOLOGIES AND SEMANTIC REPRESENTATION

Vladimir Alexiev, Data and Ontology Group, Ontotext Corp



CIDOC Congress, Dresden, Germany 2014-09-05: International Terminology Working Group: full version (HTML)

2014-09-09: Getty special session: short version (PDF)

Press O for overview, H for help.

Proudly made in plain text with reveal.js, org-reveal, org-mode and emacs.

### ONTOTEXT SCOPE OF WORK

#### http://vocab.getty.edu

- Ontology development: http://vocab.getty.edu/ontology
- Contribution to ISO 25964 ontology (latest thesauri standard)
- Complete mapping specification
- Help with R2RML conversion scripts, contrib to RDB2RDF (Perl), rrx:languageColumn extension
- GraphDB (OWLIM) repository. Enterprise Edition (clustered)
- Sem app dev (customized Forest UI), tech consulting
- SPARQL 1.1 endpoint: http://vocab.getty.edu/sparql
- Documentation (100 pages): http://vocab.getty.edu/doc
- Lots of sample queries, incl charts, geographic, etc
- Per-entity export files, explicit/total data dumps
- Help desk / support
- Presentations, scientific papers

# SEMANTIC RESOLUTION & CONTENT NEGOTIATION

All GVP, AAT and TGN URLs resolve, returning human or machine readable content through content negotiation (303 redirect). Eg about the ontology:

http://vocab.getty.edu/ontology	semantic URI, content-negotiated
http://vocab.getty.edu/ontology.html	page (application/xhtml+xml)
http://vocab.getty.edu/ontology.rdf	application/rdf+xml
http://vocab.getty.edu/ontology.ttl	text/turtle

#### Eg about an AAT subject

http://vocab.getty.edu/aat/300011154	semantic URI, con-neg
http://vocab.getty.edu/aat/300011154.html	page (application/xhtml+xml)
http://vocab.getty.edu/aat/300011154.rdf	application/rdf+xml
http://vocab.getty.edu/aat/300011154.ttl	text/turtle
http://vocab.getty.edu/aat/300011154.nt	NTriples

### **GVP VOCABULARY DATA**

#### Scope includes:

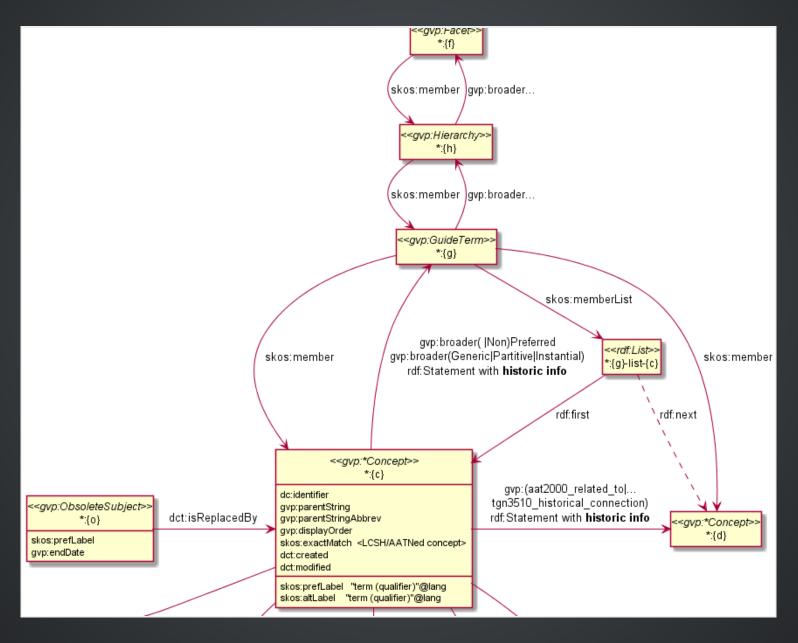
- Subjects: Concepts but also non-concepts
- Obsolete subjects (and dct:isReplacedBy)
- Terms: plain (SKOS) & rich (SKOS-XL). Term characteristics
- Languages (some custom lang tags)
- Hierarchical rels: custom & standard, distinguish BTG,BTP,BTI
- Associative rels (170 subprop of skos:related)
- Historic info on rels (rdf:Statement) and terms
- Alignment (exactMatch to LCSH)
- Sources (bibo:Document, bibo:DocumentPart with locator)
- Contributors (foaf:Agent)
- Revision history (prov:Activity)
- Thesaurus-specific (TGN place types, coordinates)

One of the richest thesauri I've seen

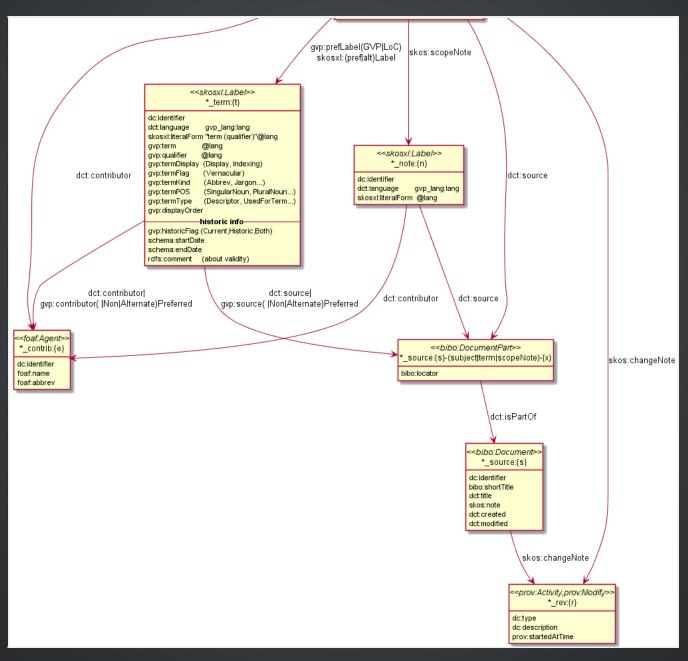
## **EXTERNAL ONTOLOGIES**

Prefix	Ontology	Used for
bibo:	Bibliography Ontology	Sources
dc:	Dublin Core Elements	common
dct:	Dublin Core Terms	common
foaf:	Friend of a Friend ontology	Contributors
iso:	ISO 25946 (latest on thesauri)	iso:ThesaurusArray, BTG/BTP/BTI
owl:	Web Ontology Language	Basic RDF representation
prov:	Provenance Ontology	Revision history
rdf:	Resource Description Framework	Basic RDF representation
rdfs:	RDF Schema	Basic RDF representation
schema:	Schema.org	common, geo (TGN)
skos:	Simple Knowledge Org System	Basic vocabulary representation
skosxl:	SKOS Extension for Labels	Rich labels
wgs:	W3C World Geodetic Survey geo	Geo (TGN)
xsd:	XML Schema Datatypes	Basic RDF representation

### GVP SEMANTIC REPRESENTATION

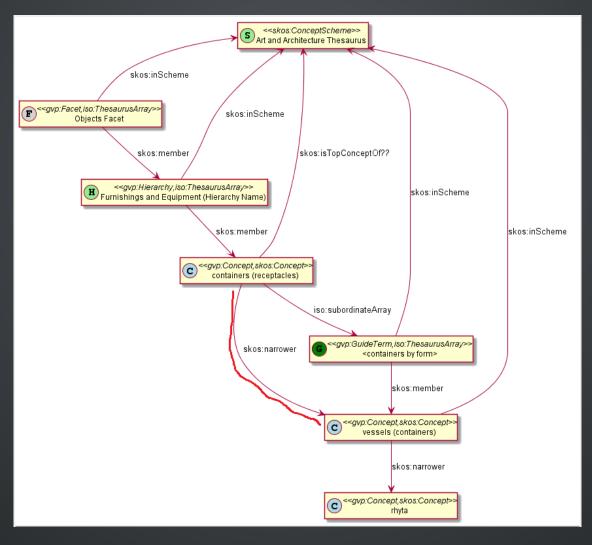


# GVP SEMANTIC REPRESENTATION (2)



### HIERARCHICAL RELATIONS

Use iso:ThesaurusArray to allow Guide Terms below Concepts. Infer cross-threading SKOS/ISO broader relations



## KEY VALUES (FLAGS) ARE IMPORTANT

Excel-driven Ontology Generation™ (getty-codes.xls to getty-codes.ttl)

Key **val** can be mapped to Custom sub-class, Custom (sub-)prop, Ontology Value (eg <term/kind/Abbreviation>)

voca	table	field	val	ObjectProperty	Class	label	domain	range	subProperty	subClass0f	ConceptSchem
	subject	record_type	F		gvp:Facet	Facet				gvp:Subject, i	so:ThesaurusArray
AAT	subject	record_type	Н		gvp:Hierarchy	Hierarchy Name				gvp:Subject, i	so:ThesaurusArray
AAT	subject	record_type	G		gvp:GuideTerm	Guide Term				gvp:Subject, i	so:ThesaurusArray
AAT	subject	record_type	С		gvp:Concept	Concept				gvp:Subject,	skos:Concept
	subject	record_type	-		gvp:ObsoleteSubject	Obsolete Subject				gvp:Subject	
TGN	subject	record_type	P		gvp:PhysPlaceConcept	Physical Place Concept				gvp:Subject,	skos:Concept
TGN	subject	record_type	A		gvp:AdminPlaceConcept	Administrative Place Con	ncept			gvp:Subject,	skos:Concept
TGN	subject	record_type	В		gvp:PhysAdminPlaceConcept	Physical and Administrat	tive Place Con	cept		gvp:Subject,	skos:Concept
	subject_rels	preferred	Р	gvp:broaderPreferred		Preferred Parent	gvp:Subject	gvp:Subject	gvp:broader		
	subject_rels	preferred	N	gvp:broaderNonPreferred		Non-Preferred Parent	gvp:Subject	gvp:Subject	gvp:broader		
	subject_rels	hier_rel_type	G	gvp:broaderGeneric		Parent (Generic)	gvp:Subject	gvp:Subject	gvp:broader		
	subject_rels	hier_rel_type	Р	gvp:broaderPartitive		Parent (Partitive)	gvp:Subject	gvp:Subject	gvp:broader		
	subject_rels	hier_rel_type	- 1	gvp:broaderInstantial		Parent (Instantial)	gvp:Subject	gvp:Subject	gvp:broader		
	term	preferred	Р	gvp:prefLabelGVP		Preferred Label for GVP		skosxl:Label			
	term	aacr2_flag	Υ	gvp:prefLabelLoC		Preferred Label for LoC	gvp:Subject	skosxl:Label			
	term	vernacular	V	gvp:termFlag		Term Flag	skosxl:Label	skos:Concept			term/flag/
	term	other_flags	Α	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	С	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	CN	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	F	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	J	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	N	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	S	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
mun. Facet a owl. Class.											

```
gvp:Facet a owl:Class;
  rdfs:isDefinedBy <http://vocab.getty.edu/ontology>;
  rdfs:subClassOf gvp:Subject, iso:ThesaurusArray;
  rdfs:label "Facet";
  rdfs:comment "One of the major divisions of a vocabulary";
  skos:example "Objects Facet (AAT), World (TGN)";
  dct:description "One of the major divisions of a vocabulary.\nExample: Objects Facet (AAT), World (TGN)".
```

### ASSOCIATIVE RELATIONS ARE VALUABLE

More Excel-driven Ontology Generation™ (assoc-rels.xls to assoc-rels.ttl)

 Relations come in owl:inverseOf pairs (or owl:SymmetricProperty self-inverse)

fcode	icode	domain (C1)	LOD frel	range (C2)		Editor frel - From C1 to C2	Editor irel - From C2 to C1	fexample	iexample
2000		any	related to	any		any - related to - any			light red (pigment) is related to gulf red
2001		,	formerly referred to	any	1 1	any - formerly referred - any		gigues formerly referred to fiddles	fiddles formerly referred to gigues
2100		,	distinguished from	any		any - distinguished from - any		distinguished from abandoned farms; naive art	abandoned farms are distinguished from historic farms; outsider art is distinguished from naive art

```
gvp:aat2000_related_to a owl:ObjectProperty;
  rdfs:subPropertyOf skos:related;
  rdfs:domain skos:Concept; rdfs:range skos:Concept;
  # domain "any"; range "any";
  dc:identifier "2000";
  skos:prefLabel "aat2000_related_to";
  dc:title "related to - any";
  skos:example "gulf red is related to light red (pigment)";
  skos:scopeNote "generic relationship, not explained";
  dct:description """any - related to - any; generic relationship, not explained.
Example: gulf red is related to light red (pigment)""" .
gvp:aat2000_related_to a owl:SymmetricProperty.
```

#### **GVP ONTOLOGY**

http://vocab.getty.edu/ontology, LOV Entry. 10 classes, 177 props: a lot are from excel, so editable by Getty



### **OBSOLETE SUBJECTS**

- AAT obsolete subjects are 4.4% of valid subjects, which shows a good rate of editorial actions
- Obsolete subjects may have been used in client data. In order not to leave such data hanging, we publish minimal information:

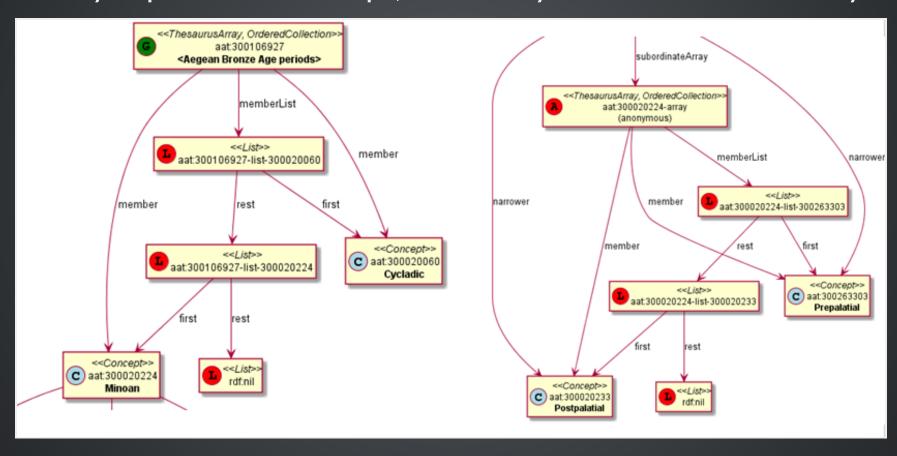
```
aat:300123456 a gvp:ObsoleteSubject; # Was made non-publishable
    skos:prefLabel "Made up subject";
    skos:inScheme aat: ;
    schema:endDate "2012-12-31T12:34:56"^^xsd:dateTime.

aat:300386746 a gvp:ObsoleteSubject; # Was merged to a dominant Subject
    skos:prefLabel "Buncheong";
    skos:inScheme aat: ;
    dct:isReplacedBy aat:300018699; # Punch'ong
    schema:endDate "2012-12-31T12:34:56"^^xsd:dateTime.
```

### USE OF ISO:THESAURUSARRAY IN GVP

Use rdf:List for ordered children.

Novelty: if parent is Concept, use anonymous ThesaurusArray



### **CONTRIBUTION TO ISO 25946**

- Contributed to ISO 25946 ontology (LOV entry)
- First industrial use of ISO 25946
- Defined appropriate combinations of BTG, BTP, BTI relations (first formally defined in ISO).

On Compositionality of ISO 25964 Hierarchical Relations (BTG, BTP, BTI), V.Alexiev, J.Lindenthal, A.Isaac. Draft paper, Presentation at NKOS 2014 Workshop, London, 12 Sep 2014

	BTGx	BTPx	BTIx
BTGx	BTGE	BTPE	no
BTPx	BTPE	BTPE	no
BTIx	BTIE	no	no

- Eg: anvil components BTP < anvils and anvil accessories > BTG
   forging and metal-shaping tools > = > BTPE
- Mount Athos BTI Orthodox religious center BTG religious center => BTIE

#### **TERMS**

#### Support multilingual labels: both SKOS (plain)...

```
aat:300198841 a skos:Concept , gvp:Subject , gvp:Concept ;
   skos:prefLabel "rhyta"@el-latn , "rhyta"@en , "rhytons"@es , "rhytons"@fr ;
   skos:altLabel "rhyta"@es , "rhyton"@es , "rhyton"@en , "rhyton"@el-latn ...;
   skosxl:prefLabel aat_term:1000198841-en , aat_term:1000198841-el-Latn ...;
   skosxl:altLabel aat_term:1000198841-es , aat_term:1000297235-en ...
```

#### ... and rich info in SKOS-XL:

```
aat term:1000198841-en a skosxl:Label;
 dc:identifier "1000198841";
 dct:language aat:300388277 , gvp lang:en ; # owl:sameAs
 dct:contributor aat contrib:10000000 , aat contrib:10000131 ;
 skosxl:literalForm "rhyta"@en ; #### with Qualifier if applicable
                                          #### no qualifier
 gvp:term "rhyta"@en ;
 qvp:displayOrder "1"^^xsd:positiveInteger ;
 gvp:termType <term/type/Descriptor> ; #### Descr/AltDescr/UseFor
 gvp:termPOS <term/POS/PluralNoun>; #### Part of Speech
 gvp:contributorPreferred aat contrib:10000000 , aat contrib:10000088 ;
 gvp:contributorNonPreferred aat contrib:10000131;
 gvp:sourcePreferred aat source:2000051089-term-1000198841 ...;
 dct:source aat source:2000024811 , aat source:2000052946 ...;
 gvp:sourceNonPreferred aat source:2000052946;
 gvp:sourceAlternatePreferred aat source:2000048328-term-1000198841 .
```

### SOME CUSTOM LANGUAGE TAGS

Despite the richness of IANA tags (9000), we had to define new tags, using several extension mechanisms:

- Private language, e.g.
  - x-byzantin-Latn: Byzantine Greek (transliterated)
  - x-frisian (IANA/ISO has codes for predecessor Old Frisian and dialects West, Saterland and North Frisian)
- Private language used in specific region, e.g.
  - qqq-002: African language (not specified which)
  - qqq-ET: Ethiopian (not specified: Boro/Borna, Karo...)
- Private modifier, e.g.
  - grc-Latn-x-liturgic: Liturgical Greek
  - ber-Latn-x-dialect: Berber Dialects (transliterated)
  - fa-Latn-x-middle: Persian, Middle (transliterated)
  - zh-Latn-pinyin- x-notone: transliterated Pinyin without tones

### SOURCES

#### bibo:Document or bibo:DocumentPart

```
aat_source:2000051089 a bibo:Document;
   dc:identifier "2000051089"
   bibo:shortTitle "AATA database (2002-)";
   dct:title "Getty Conservation Institute (GCI). database of AATA Online...".
aat_source:2000051089-term-1000198841 a bibo:DocumentPart;
   dct:isPartOf aat_source:2000051089;
   bibo:locator "128257 checked 26 January 2012".
```

#### Applied to subject, term, scopeNote:

```
aat:300198841 # subject (rhyta)
  dct:source aat_source:2000030301-subject-300198841;
  dct:source aat_source:2000052378.

aat_term:1000198841-en # term "rhyta"@en
  gvp:sourceNonPreferred aat_source:2000049728;
  dct:source aat_source:2000051089-term-1000198841.

aat_scopeNote:34904 # scopeNote
  dct:source aat_source:2000046502.
```

### CONTRIBUTORS

#### foaf:Agent

```
aat_contrib:10000131 a foaf:Agent;
  dc:identifier "10000131";
  foaf:nick "CDBP-DIBAM";
  foaf:name "Centro de Documentación de Bienes Patrimoniales...".
```

#### Applied to subject, term, scopeNote:

```
aat:300198841 # subject "rhyta"
  dct:contributor aat_contrib:10000131;
  dct:contributor aat_contrib:10000000.

aat_term:1000198841-en # term "rhyta"@en
  gvp:contributorNonPreferred aat_contrib:10000131;
  gvp:contributorPreferred aat_contrib:10000000.

aat_scopeNote:34904 # scopeNote
  dct:contributor aat_contrib:10000000.
```

### HISTORIC INFO

Includes dates of applicability, historicFlag, comment. Applied to terms; hier & assoc rels, place types (using rdf:Statement)

```
aat term:1000002693-en a skosx1:Label;
 skosxl:literalForm "lambruscatura"@en ;
 qvp:historicFlag <http://vocab.getty.edu/historic/historic>;
 schema:startDate "0900"^^xsd:gYear ;
 schema:endDate "1700"^^xsd:gYear;
 rdfs:comment "Medieval term for wainscoting".
aat rel:300020271-aat2812 followed-300020269 a rdf:Statement;
 rdf:subject
                  aat:300020271;
                                       # Second Dynasty (Egyptian)
 rdf:predicate
                  gvp:aat2812 followed;
 rdf:object
                  aat:300020269;
                                       # First Dynasty (Egyptian)
 rdfs:comment
                 "Second Dynasty began ca. 2775 BCE";
 schema:startDate "-2785"^^xsd:gYear;
 schema:endDate "-2765"^^xsd:gYear.
tgn:7011179-placeType-300008347 a rdf:Statement;
 rdf:subject
                  tgn:7011179;
                                        # Siena
 rdf:predicate
                  gvp:placeTypePreferred;
 rdf:object
                  aat:300008347; # inhabited place
 rdfs:comment
                 "settled by Etruscans (flourished 6th century BCE)";
 schema:startDate "-0800"^^xsd:qYear;
 gvp:displayOrder "1"^^xsd:positiveInteger.
```

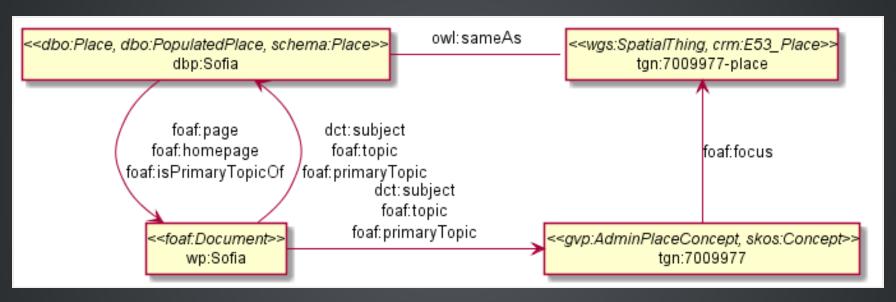
### **REVISION HISTORY**

#### PROV is too complex, so we simplify:

```
aat:300018699
 skos:changeNote aat rev:12345, aat rev:12346, aat rev:12347;
 prov:wasGeneratedBy aat rev:12345;
 dct:created "2014-01-02T01:02:03"^^xsd:dateTime;
 dct:modified "2014-01-03T01:02:03"^^xsd:dateTime;
 dct:issued "2014-01-04T01:02:03"^^xsd:dateTime.
aat rev:12345 a prov:Activity, prov:Create;
 dc:type "created";
 prov:startedAtTime "2014-01-02T01:02:03"^^xsd:dateTime.
aat rev:12346 a prov:Activity, prov:Modify;
 prov:used aat:300018699;
 dc:type "term added";
 dc:description "leggings, puttee (1000248060)";
 prov:startedAtTime "2014-01-03T01:02:03"^^xsd:dateTime.
aat rev:12347 a prov:Activity, prov:Publish;
 prov:used aat:300018699;
 dc:type "issued";
 prov:startedAtTime "2014-01-04T01:02:03"^^xsd:dateTime.
```

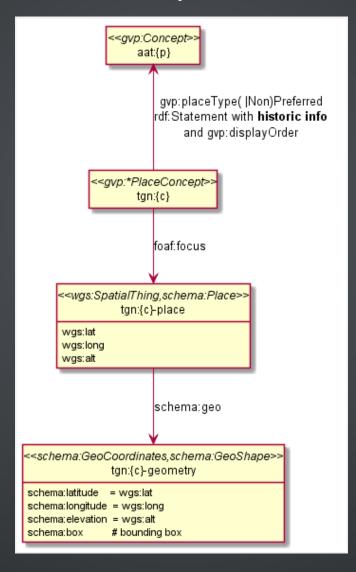
### TGN SPECIFICS: CONCEPT-PLACE DUALITY

Duality between Concept and its denotation (ala VIAF, UK BL, FR BnF, SE KB...)

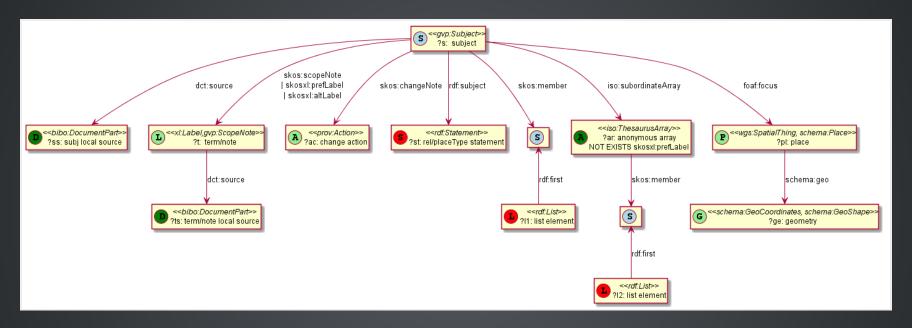


### TGN SEMANTIC REPRESENTATION

Place types (TGN->AAT), Concept-Place duality, coordinates



### CONSTRUCT QUERY: GET & CACHE ALL DATA FOR SUBJECT



- All data for these subsidiary objects is served by the resource URL
- Cached, thus served quickly
- Served in RDF/XML, N3/Turtle, NTriples, JSON, soon JSON-LD

### DOCUMENTATION

#### Getty Vocabularies: Linked Open Data

#### Semantic Representation

Version: 2.0

Last updated: 19 Aug 2014

HTML version: http://vocab.getty.edu/doc/ (for link http://vocab.getty.edu/doc/gvp-lod.j PDF version: http://www.getty.edu/research/tools Formerly at:

Initial version: Vladimir Alexiev, Joan Cobb, Greg

Updates: Vladimir Alexiev, Joan Cobb

#### Table of Contents

#### Introduction

The Getty Vocabularies and LOD

About the AAT 1.1.1

1.1.2 About the TGN

Revisions, Review, Feedback

1.2.1 Revisions

> 1.2.1.1 Version 1.0

> 1.2.1.2 Version 1.1

1.2.1.3 Version 1.2

1.2.1.4 Version 1.3

1.2.1.5 Version 2.0

1.2.1.6 Future Versions

External Review Process

Providing Feedback

124 Disclaimer

Abbreviations

RDF Turtle

Prefixes

1.5.1 External Prefixes

Descriptive Prefixes

GVP URLs and Prefixes 1.6

Common GVP URLs

162 AAT URLs

1.6.3 TGN URLs 1.6.4 Using GVP URLs

1.6.5 Named Graphs

Semantic Resolution

External Ontologies 1.8.1 DC and DCT

1.8.2 SKOS and SKOS-XL

1.8.3 ISO 25964 1.8.4 BIBO

1.8.5 FOAF

1.8.6 PROV

> 1.8.6.1 dct:modified

1.8.6.2 dct:creator+dct:created

Geographic Ontologies

1.8.7.1 W3C WGS Geo Ontology

1.8.7.2 Schema.org Geographic Features

GVP Ontology

Semantic Representation

Semantic Overview

Subject

2.2.1 Subject Types

Subject Hierarchy

2.3.1Standard Hierarchical Relations

**GVP Hierarchical Relations** 2.3.2

Hierarchy Structure 2.3.3

2.3.4 Top Concepts

Sort Order

Sorting with Thesaurus Array

skos:member Structure 2.4.1.1

2.4.1.2 skos:memberList Structure 2.4.1.3 Full Representation

Very detailed: 100 pages! Linkable anchors:

vocab.getty.edu/doc/#Full Text Search

Associative Relationships

Relationships Table 2.5.1

Relationship Cross-Walk

Relationship Representation

Obsolete Subject 2.6

Language

IANA Language Tags 2.7.1

GVP Language Tags 2.7.2

Language Tag Case

Language Tags and Sources

Language Dual URLs

Term

2.8.1 Term Characteristics

Importance of the Vernacular Flag

Scope Note

2.10 Identifiers

Notations 2.11

Source 2.12

> Local Sources 2.12.1

Contributor

Historic Information

2.14.1Applying to Terms

2.14.2 Applying to Relations and Place Types

2.15 Revision History

Revision History Representation 2.15.1

2.15.2 Revision History for Subject

Revision History for Source 2.15.3

TGN Specifics

TGN Overview

TGN Place Types

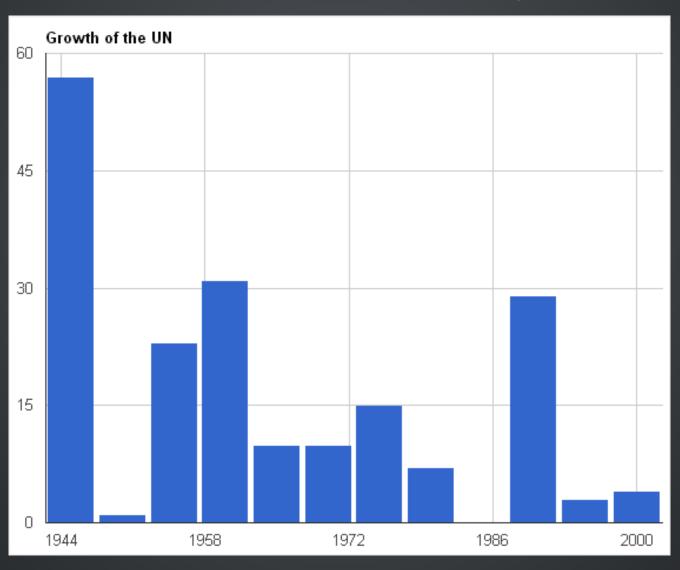
Concept vs Place Duality

Cons of the Dual Approach 3.3.1

Co-reference and Co-denotation 3.3.2

### SAMPLE QUERY: BAR CHART WITH SPARQL

Number of UN members per year. See doc or jsfiddle with it



# THANKS FOR YOUR TIME!

mailto:vladimir.alexiev@ontotext.com

