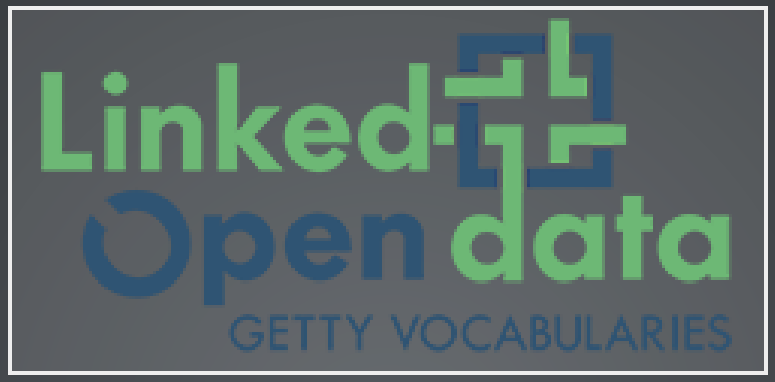
GVP LOD: ONTOLOGIES AND SEMANTIC REPRESENTATION

Vladimir Alexiev, Data and Ontology Group, Ontotext Corp



CIDOC Congress, Dresden, Germany
2014-09-05: International Terminology Working Group:
full version (HTML, PDF, slideshare)
2014-09-09: Getty special session:
short version (HTML, PDF, slideshare)

Press O for overview, H for help.
Proudly made in plain text with reveal.js, org-reveal, org-mode and emacs.

TABLE OF CONTENTS

GETTY VOCABULARIES LOD

http://vocab.getty.edu

- Art and Architecture Thesaurus (AAT): released Feb 2014
- Thesaurus of Geographic Names (TGN): released Aug 2014

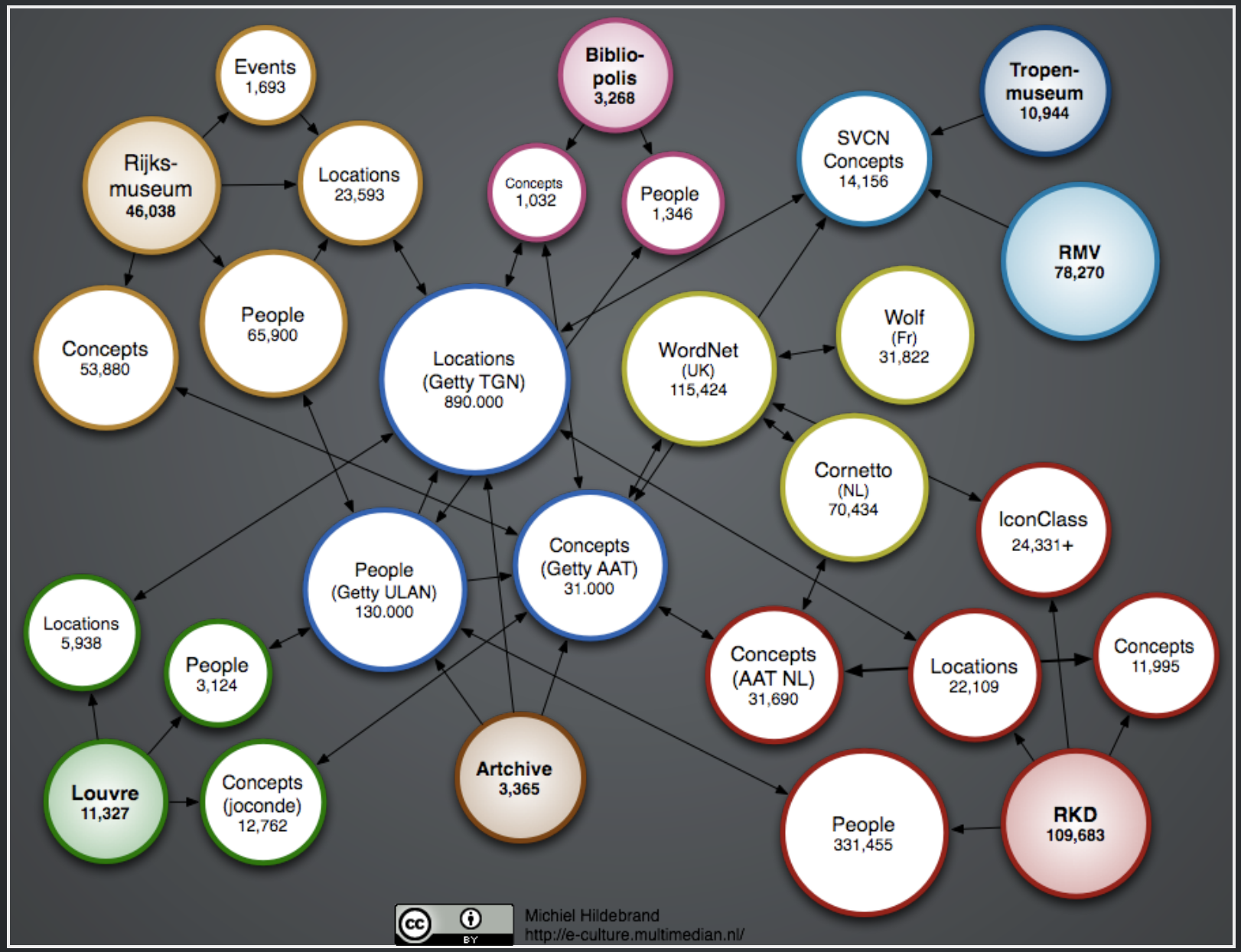
Work continues with:

- Unified List of Artist Names (ULAN)
- Cultural Object Names Authority (CONA)
- Getty Museum data
- AATA bibliography

Museum and CONA are more complex, involves LIDO/CDWA-lite XML to CIDOC CRM (RDF)

CULTURAL HERITAGE LOD

Working at the center. (Shows thesauri only, not yet CONA/Museum data)

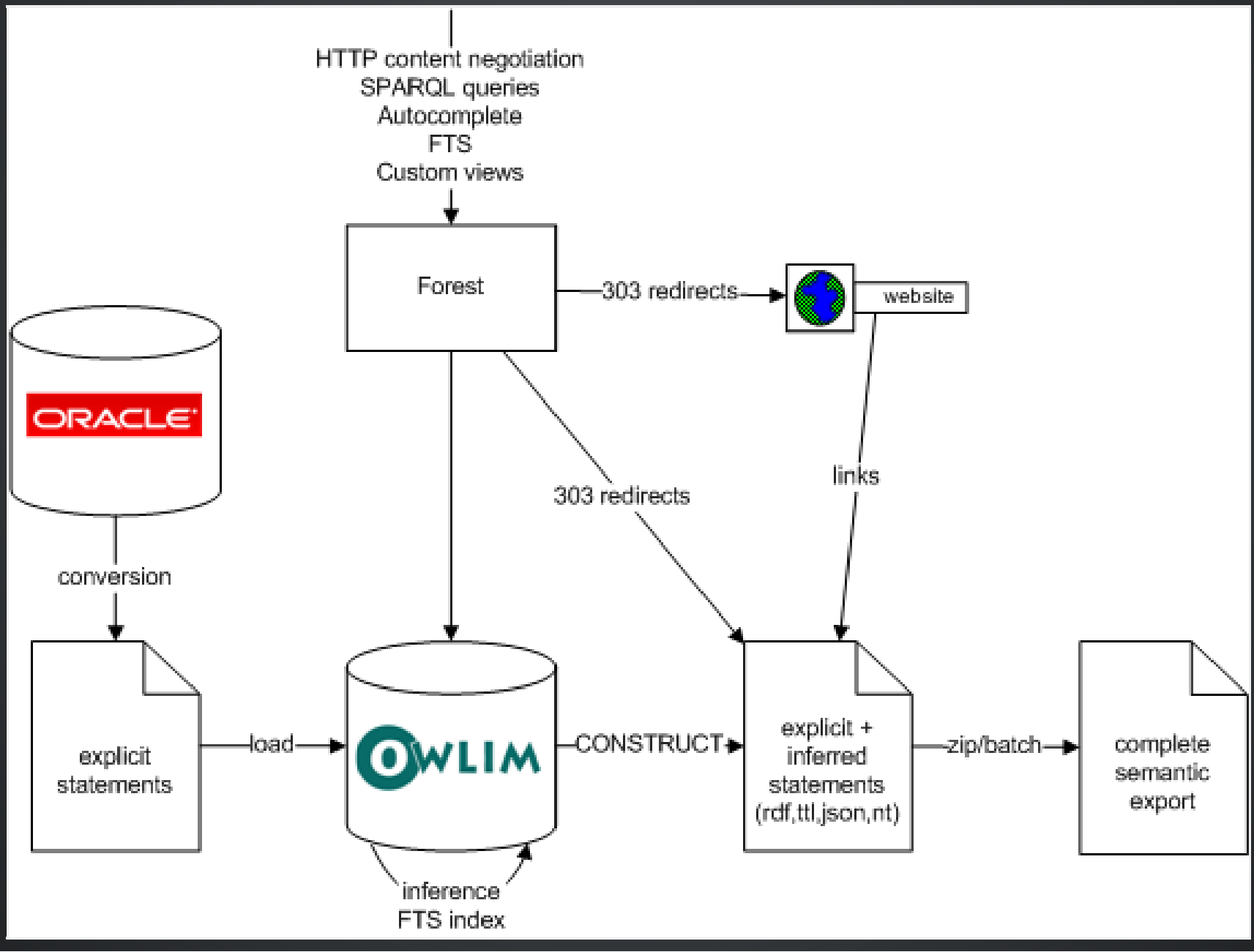


ONTOTEXT SCOPE OF WORK

- Semantic/ontology development: http://vocab.getty.edu/ontology
- Contributed to ISO 25964 ontology (latest standard on thesauri). Provided implementation experience, suggestions and fixes
- Complete mapping specification
- Help implement R2RML scripts working off Getty's Oracle database, contribution to Perl implementation (RDB2RDF), R2RML extension (rrx:languageColumn)
- Work with a wide External Reviewers group (people from OCLC, Europeana, ISO 25964 working group, etc)
- GraphDB (OWLIM) semantic repository. Enterprise Edition (clustered for high-availability)
- Semantic application development (customized Forest user interface) and tech consulting
- SPARQL 1.1 compliant endpoint: http://vocab.getty.edu/sparql
- Comprehensive documentation (100 pages): http://vocab.getty.edu/doc
- Lots of sample queries, including charts, geographic queries, etc
- Per-entity export files, explicit/total data dumps. Many formats: RDF, Turtle, NTriples, JSON, JSON-LD (upcoming)
- Help desk / support
- Presentations, scientific papers

GVP LOD ARCHITECTURE

Quite straightforward



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	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, content-negotiated
http://vocab.getty.edu/aat/300011154.html	Forest 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
http://vocab.getty.edu/aat/300011154.json	JSON (to change to .rj)
http://vocab.getty.edu/aat/300011154.jsonld	JSON-LD (upcoming)

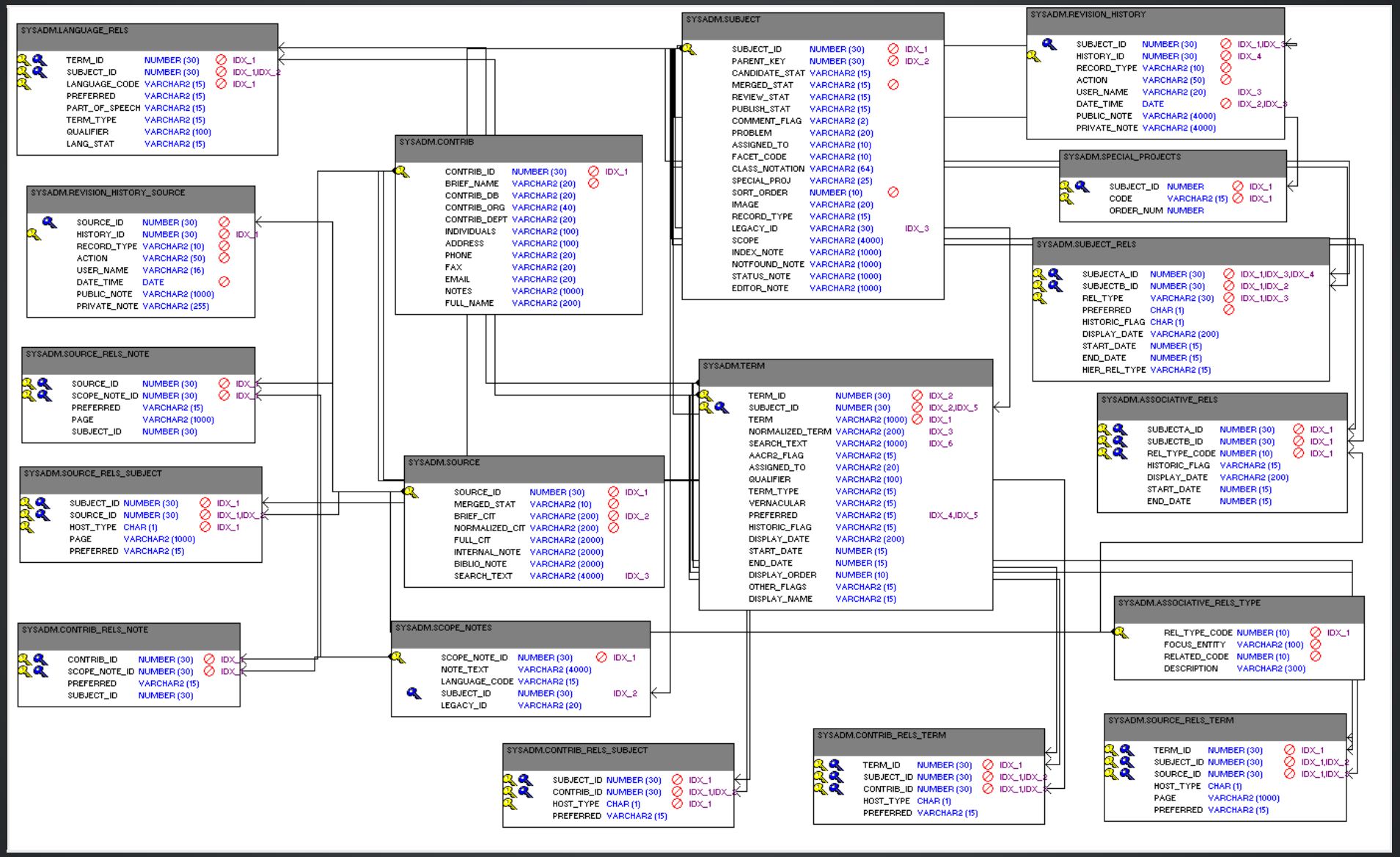
GVP VOCABULARY DATA

Scope includes:

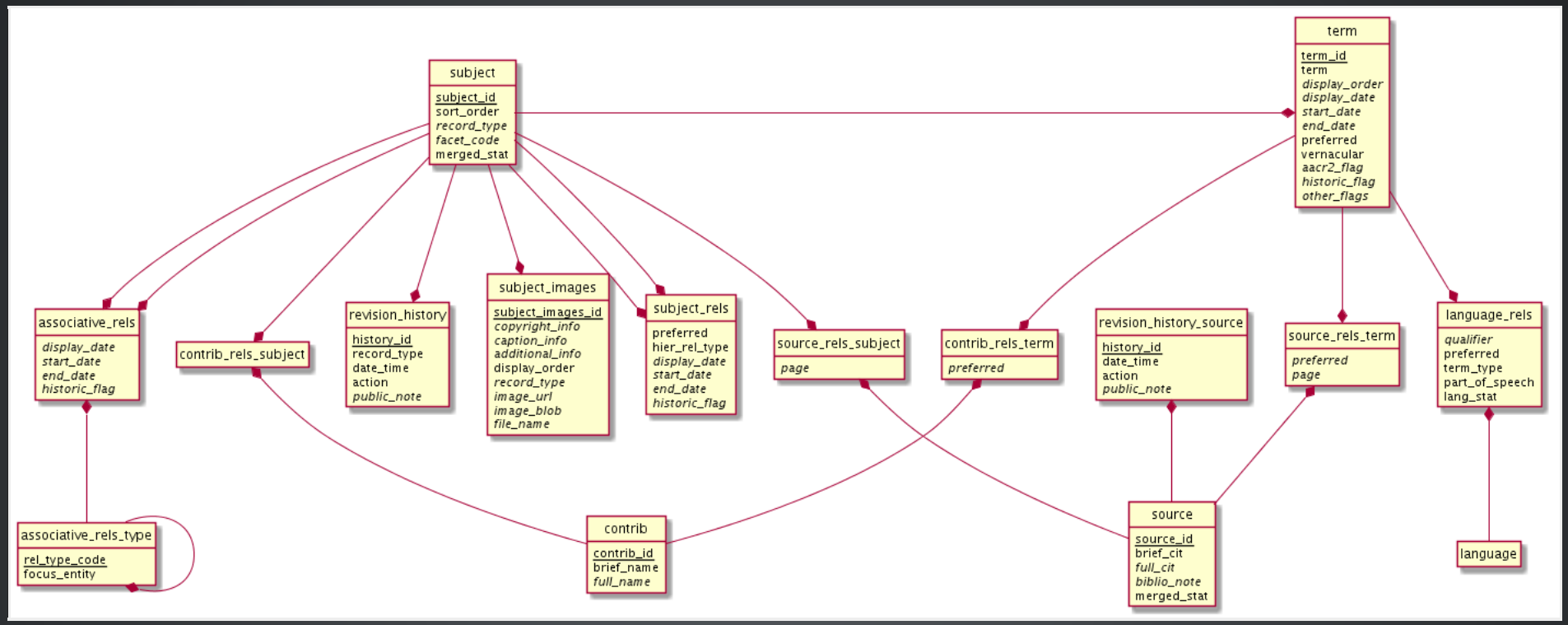
- Subjects: Concepts but also non-concepts
- Terms as plain (SKOS) and rich (SKOS-XL) labels. Term characteristics
- Hierarchical relations: custom & standard, distinguish BTG,BTP,BTI
- Associative Relations
- Historic info on rels (rdf:Statement) and terms
- Obsolete subjects
- Alignment (exactMatch to LCSH)
- Sources (bibo:Document, bibo:DocumentPart with locator)
- Contributors (foaf:Agent)
- Revision history (prov:Activity)
- Thesaurus-specific data (for now: TGN place types, coordinates

Richer than any other SKOS thesaurus I've seen

AAT RELATIONAL SCHEMA



AAT CONCEPTUAL DIAGRAM



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 Organization System	Basis vocabulary representation
skosxl:	SKOS Extension for Labels	Rich labels
wgs:	W3C World Geodetic Survey geo	Geo (TGN)
xsd:	XML Schema Datatypes	Basic RDF representation

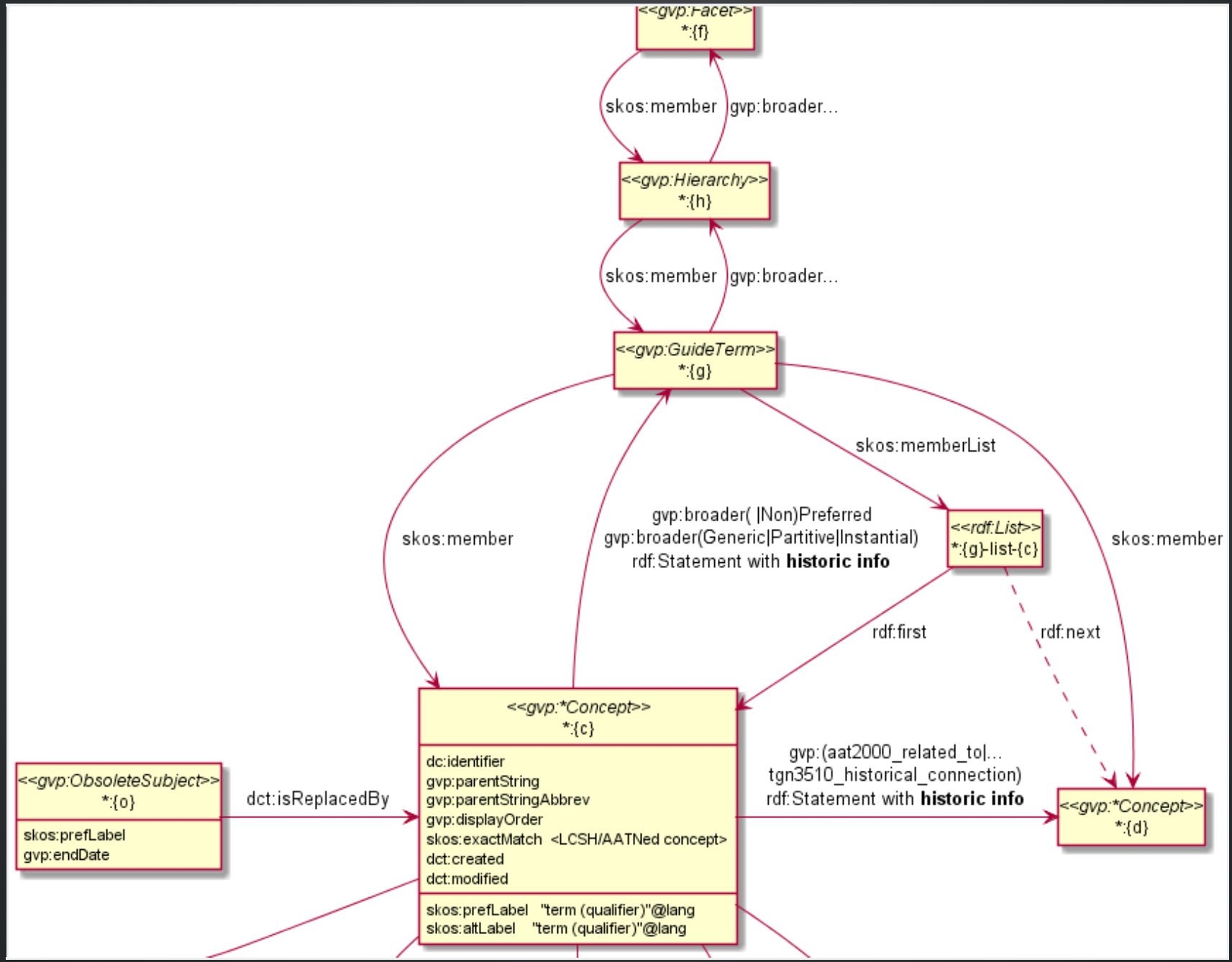
AUXILIARY ONTOLOGIES

Prefix	Ontology	Used for
luc:	OWLIM's built-in Lucene	Full Text index & queries
ontogeo:	OWLIM geo-spatial extensions	Geo-spatial index & queries
ptop:	Ontotext PROTON top-level ontology	Inferencing (Extended Property Constructs)
rr:	Relational to RDF Mapping Language	Conversion Oracle->RDF
rrx:	R2RML extension	rrx:languageColumn

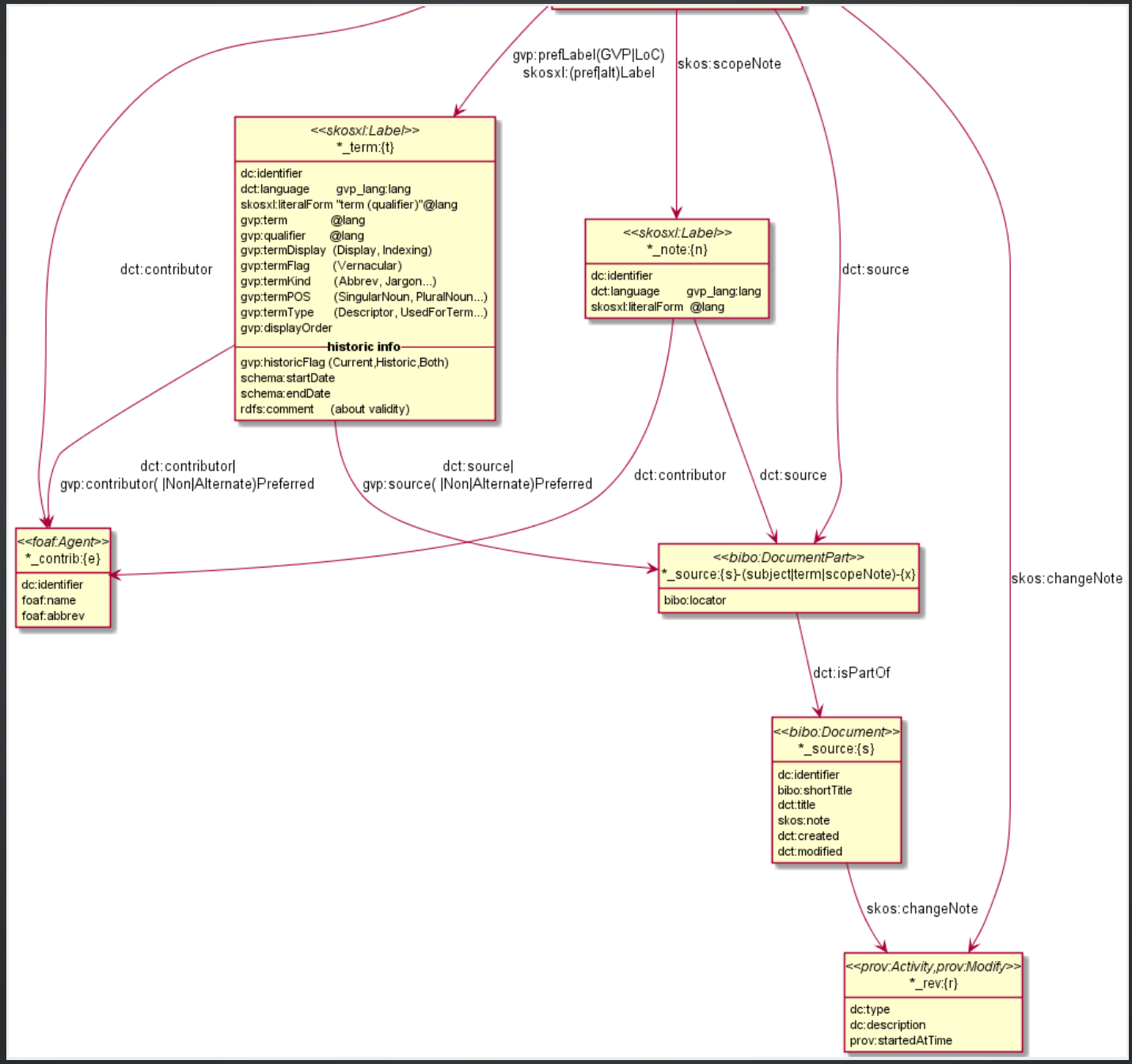
DESCRIPTIVE INFO ONTOLOGIES

Prefix	Ontology	Used for
adms:	Asset Description Metadata Schema	Dataset description
cc:	Creative Commons Rights Expressions	License rights
dcat:	Data Catalog Vocabulary	Dataset description
dctype:	DCMI Type Vocabulary	Dataset class
fmt:	RDF formats used in datasets	Formats of data dumps
sd:	SPARQL Service Description	SPARQL endpoint capabilities (future)
vaem:	Vocabulary Attaching Essential Metadata	Not used yet
vann:	Vocabulary for annotating vocabularies	Namespace and prefix
vcard:	vCard (contact info)	Contact info
vdpp:	Vocabulary for Dataset Publ Projects	Not used yet
voaf:	Vocabulary of a Friend	Linked Open Vocabularies (LOV)
voag:	Vocabulary Of Attribution and Governance	Frequency of publication
void:	Vocabulary of Interlinked Datasets	Basis descr, LOD registration
wdrs:	Protocol for Web Description Resources	Described by from dataset to doc
wv:	A vocabulary for waivers of rights	License rights

GVP SEMANTIC REPRESENTATION



GVP SEMANTIC REPRESENTATION (2)



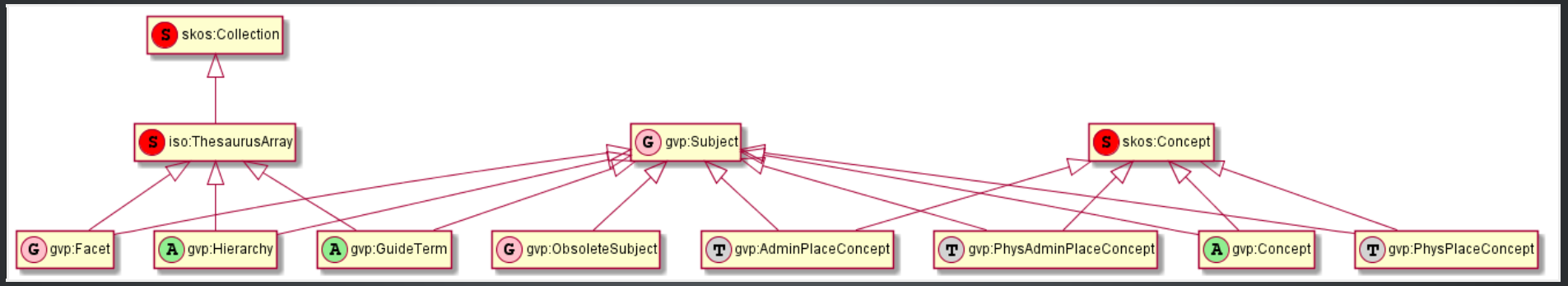
GVP SUBJECT CLASSES

- GVP Subjects include both Concepts and non-concepts (for organizing the hierarchy, not for indexing)

We handle "impedance mismatch" with

- SKOS: restrict skos:related, infer skos:broader
- ISO: infer iso:broaderGeneric/Partitive/Instantial

S=Standard, G=GVP common, A=AAT, T=TGN



OBSOLETE SUBJECTS

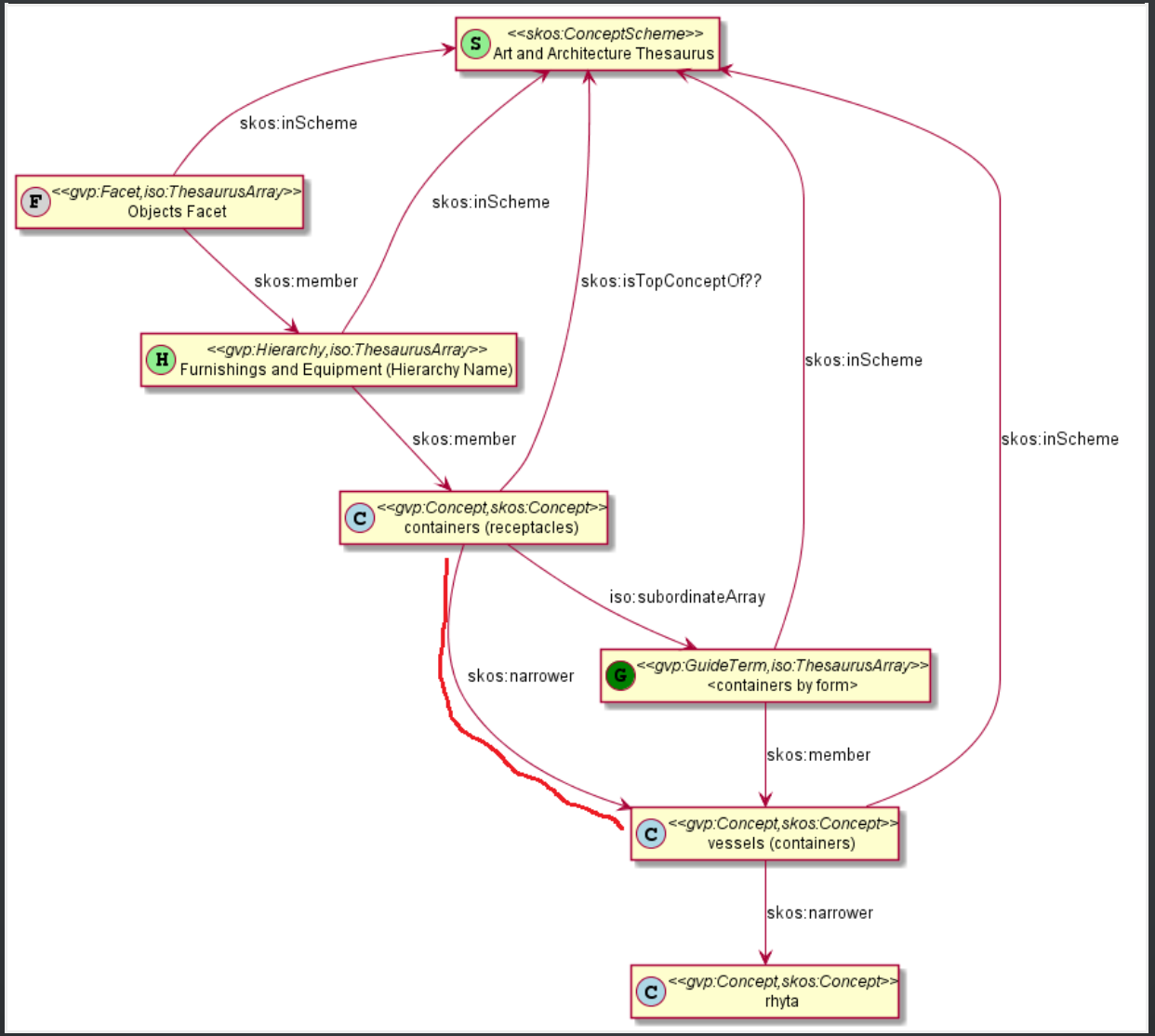
- AAT obsolete subjects are 4.4% of valid subjects, which shows a good rate of editorial actions
- Obsolete subjects may have been used in client data. In order not to leave such data hanging, we publish minimal information:

```
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.
```

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	subClassOf	ConceptSchem
	subject	record_type	F		gvp:Facet	Facet				gvp:Subject, iso:ThesaurusArray	
AAT	subject	record_type	H		gvp:Hierarchy	Hierarchy Name				gvp:Subject, iso:ThesaurusArray	
AAT	subject	record_type	G		gvp:GuideTerm	Guide Term				gvp:Subject, iso:ThesaurusArray	
AAT	subject	record_type	C		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 Concept				gvp:Subject, skos:Concept	
TGN	subject	record_type	B		gvp:PhysAdminPlaceConcept	Physical and Administrative Place Concept				gvp:Subject, skos:Concept	
	subject_rels	preferred	P	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	P	gvp:broaderPartitive		Parent (Partitive)	gvp:Subject	gvp:Subject	gvp:broader		
	subject_rels	hier_rel_type	I	gvp:broaderInstantial		Parent (Instantial)	gvp:Subject	gvp:Subject	gvp:broader		
	term	preferred	P	gvp:prefLabelGVP		Preferred Label for GVP	gvp:Subject	skosxl:Label			
	term	aacr2_flag	Y	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	A	gvp:termKind		Term Kind	skosxl:Label	skos:Concept			term/kind/
AAT	term	other_flags	C	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/

```
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)
- Should we make a subproperty hierarchy?

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

http://vocab.getty.edu/ontology, LOV Entry (10 classes, 177 props)

Table of contents

Ontologies (1):

Getty Vocabulary Program ontology

Classes (10):

AdminPlaceConcept , Concept , Facet , GuideTerm , Hierarchy , ObsoleteSubject , PhysAdminPlaceConcept , PhysPlaceConcept , Scope Note , Subject

Properties (177):

aat2000_related_to , aat2001_formerly_referred_to , aat2100_distinguished_from , aat2110_meaning-usage_overlaps_with , aat2203_associated_with , aat2205_causes-is_required , aat2206_caused_by-requires , aat2208_locus-setting_for , aat2209_used-located_in , aat2211_produce , aat2212_produced_by , aat2215_required_for , aat2216_require , aat2218_used-function_as , aat2219_have_form , aat2222_act_upon , aat2223_are_acted_upon , aat2281_have_counterpart , aat2285_practiced-studied_by , aat2286_practice-study , aat2291_locus-setting_for , aat2292_work-live_in , aat2294_locus-setting_for , aat2295_located_in , aat2311_performed_by , aat2312_perform , aat2315_used_by , aat2316_use , aat2318_involved_in , aat2319_involves , aat2321_used_by , aat2322_use , aat2325_created_by , aat2326_create , aat2328_involved_with , aat2329_involves , aat2332_affiliated_with , aat2333_have_affiliates , aat2335_associated_with , aat2336_has_associates , aat2397_focus_of , aat2398_focuses_on , aat2408_locus-setting_for , aat2409_takes_place_in , aat2411_involved_in , aat2412_involves , aat2415_required_for , aat2416_requires , aat2418_uses , aat2419_used_for , aat2421_locus-setting_for , aat2422_takes_place_in , aat2424_produced_by , aat2425_produces , aat2427_produced_by , aat2428_produces , aat2431_required_for , aat2432_requires , aat2434_contextualized_in , aat2435_context_for , aat2501_made_of-require , aat2502_material_for , aat2504_used-located_in , aat2505_locus-setting_for , aat2507_produce-process , aat2508_produced-processed_by , aat2551_reflect-produced_by , aat2552_reflected_in-produces , aat2554_reflects , aat2555_reflected_in , aat2557_reflects , aat2558_reflected_in , aat2562_locus-setting_for , aat2563_located_in , aat2601_possessed_by-existing_in , aat2602_possess-exist_in , aat2604_possessed_by-existing_in , aat2605_possess-exist_in , aat2607_caused_by-requires , aat2608_causes-required_for , aat2612_possessed_by-existing_in , aat2613_possess-exist_in , aat2801_conjuncted_with , aat2802_exemplified_by , aat2803_example_of , aat2805_contextualized_in , aat2806_provide_context , aat2807_derived_from_common_source , aat2809_coexisted_with , aat2811_preceded , aat2812_followed , aat2814_constituent_of , aat2815_composed_of , aat2817_derived-made_from , aat2818_source_for , aat2821_based_on , aat2822_basis_of , aat2824_has_parallels_with , aat2826_used_with , aat2828_use-require , aat2829_used-required_for , aat2831_associated_with , aat2833_ancestor_of , aat2834_decendant_of , aat2836_derived-made_from , aat2837_source_for , aat2841_derived-made_from , aat2842_source_for , aat2845_used-located_in , aat2846_locus-setting_for , aat2848_involved_in , aat2849_involves , aat2852_involved_with , aat2853_involve , aat2875_characteristic_of , aat2876_characterized_by , aat2878_preceded-source_for , aat2879_followed-developed_from , aat2881_reflected_in , aat2882_reflect , aat2884_involved_with , aat2885_involves , aat2891_exemplified_by , aat2892_example_of , aat2894_exemplified_by , aat2895_example_of , aat2900_miscellaneous_relationship , broader , broaderExtended , broaderGeneric , broaderGenericExtended , broaderInstantial , broaderInstantialExtended , broaderNonPreferred , broaderPartitive , broaderPartitiveExtended , broaderPreferred , broaderPreferredExtended , contributorAlternatePreferred , contributorNonPreferred , contributorPreferred , displayOrder , historicFlag , narrower , narrowerExtended , parentString , parentStringAbbrev , placeTypeNonPreferred , placeTypePreferred , prefLabelGVP , prefLabelLoC , qualifier , sourceAlternatePreferred , sourceNonPreferred , sourcePreferred , term , termDisplay , termFlag , termKind , termPOS , termType , tgn3000_related_to , tgn3001_distinguished_from , tgn3005_possibly_identified_as , tgn3101_adjacent_to , tgn3102_coextensive_with , tgn3110_meaning-usage_overlaps_with , tgn3201_capital_of , tgn3202_capital_is , tgn3301_ally_of , tgn3317_member_of , tgn3318_member_is , tgn3401_moved_from , tgn3402_moved_to , tgn3411_successor_of , tgn3412_predecessor_of , tgn3510_historical_connection

Individuals (44):

Abbreviation , Adjectival , Alternate Descriptor , Art and Architecture Thesaurus , Both , Both Singular and Plural , Chemical Name , Common term , Current , Descriptor , FIPS Code , Full term , Getty Research Institute , Historic , Historic Flag concept scheme , ISO alpha-2 code , ISO alpha-3 code , ISO numeric-2 code , ISO numeric-3 code , Jargon or Slang , Loan Term , Neologism , Noun , Official Name , Ontotext Corp , Part of Speech concept scheme , Past Participle , Plural Noun , Provisional Name , Pseudonym , Scientific or Technical term , Singular Noun , Site Name , Term Display concept scheme , Term Flag concept scheme , Term Kind concept scheme , Term Type concept scheme , Thesaurus of Geographic Names , US Postal Service Code , Use for Display , Use in Indexes/lists , Used for Term , Verbal Noun/Gerund , Vernacular

GVP ONTOLOGY: A CLASS

C Subject (OWL Class)

Definition | In use

Node in a GVP vocabulary hierarchy.
May be Facet, GuideTerm, Hierarchy, Concept, AdminPlaceConcept, PhysPlaceConcept, PhysAdminPlaceConcept, or ObsoleteSubject.
Implemented as skos:Concept or iso:ThesaurusArray and skos:Collection

Definition

The **URI** of this class is `http://vocab.getty.edu/ontology#Subject`

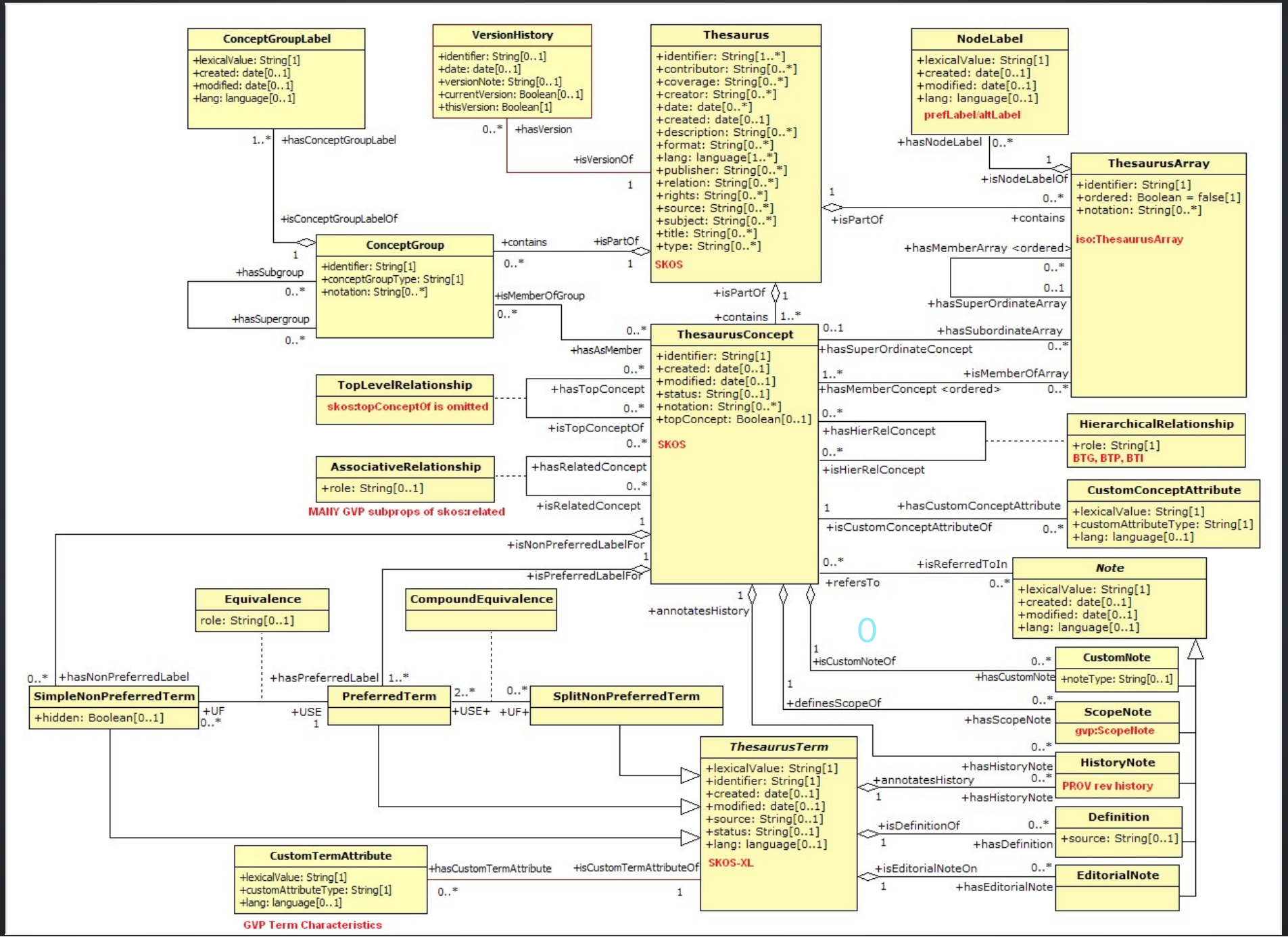
This class is a super class of C AdminPlaceConcept C Concept C Facet C GuideTerm C Hierarchy C ObsoleteSubject C PhysAdminPlaceConcept C PhysPlaceConcept

In use

This class is defined by Getty Vocabulary Program ontology

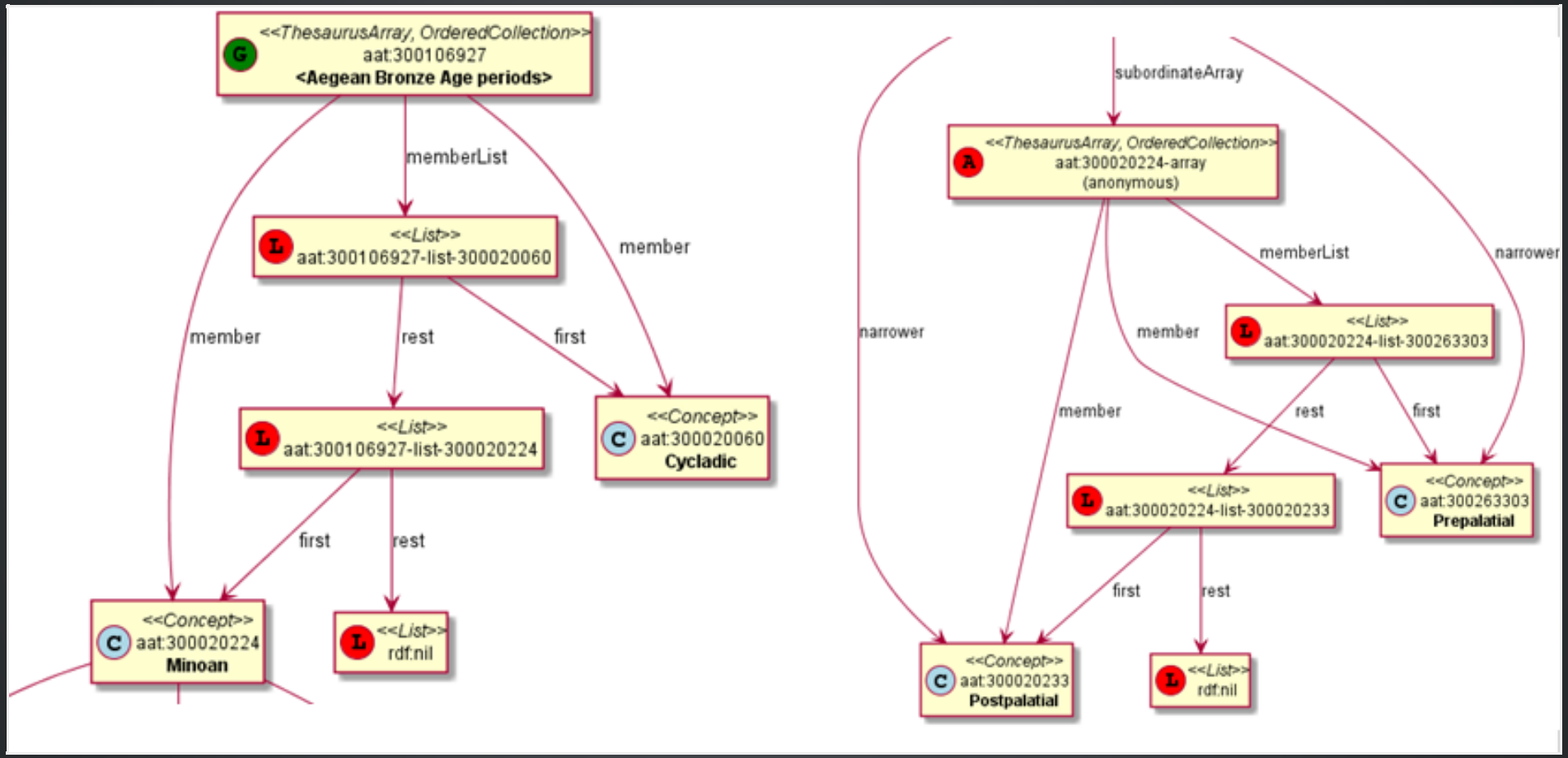
This class is used in P broader P broaderExtended P broaderGeneric P broaderGenericExtended P broaderInstantial P broaderInstantialExtended P broaderNonPreferred P broaderPartitive P broaderPartitiveExtended P broaderPreferred P broaderPreferredExtended P narrower P narrowerExtended P parentString P parentStringAbbrev P placeTypeNonPreferred P placeTypePreferred P prefLabelGVP P prefLabelLoC

ISO 25946: LATEST STANDARD ON THESAURI



USE OF ISO:THESAURUSARRAY IN GVP

Use for ordered children. Novelty: if parent is Concept, use anonymous array. Careful crafting of URLs to make rdf:List



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 at DL 2014, London, 12 Sep 2014

	BTGx	**BTPx**	**BTIx**
BTGx	BTGE: numerous examples	BTPE: beak irons BTG anvil components BTP <anvils and anvil accessories>	no
BTPx	BTPE: anvil components BTP <anvils and anvil accessories> BTG <forging and metal-shaping tools>	BTPE: Sofia BTP Bulgaria BTP Europe	no: Sofia BTP Bulgaria BTI country, but Sofia is no country
BTIx	BTIE: Mt Athos BTI orthodox religious center BTG Christian religious center	no	no

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 , "rytons"@nl ;
  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 , aat_contrib:10000088 ;
  skosxl:literalForm "rhyta"@en ;                #### with Qualifier if applicable
  gvp:term "rhyta"@en ;                          #### no qualifier
  gvp:displayOrder "1"^^xsd:positiveInteger ;
  gvp:termType <http://vocab.getty.edu/aat/term/type/Descriptor> ; #### Descr/AltDescr/UseFor
  gvp:termPOS <http://vocab.getty.edu/aat/term/POS/PluralNoun> ;   #### Part of Speech
  gvp:contributorPreferred aat_contrib:10000000 , aat_contrib:10000088 ;
  gvp:contributorNonPreferred aat_contrib:10000131 ;
  gvp:sourcePreferred aat_source:2000024811 , aat_source:2000051089-term-1000198841...;
  dct:source aat_source:2000024811 , aat_source:2000052946 , aat_source:2000049728...;
  gvp:sourceNonPreferred aat_source:2000052946 ;
  gvp:sourceAlternatePreferred aat_source:2000048328-term-1000198841 .
```

LANGUAGES

IANA Language Subtag Registry: 9000 registrations (broken down by Type and Scope):

- 7769 languages
- 227 extlangs, e.g. ar-auz (Uzbeki Arabic)
- 116 language collections, e.g. bh (Bihari languages)
- 62 macrolanguages, e.g. zh (Chinese), cr (Cree)
- 4 special languages, e.g. und (Undetermined)
- 162 scripts, eg Latn (Latin), Japn (Japanese)
- 301 regions, e.g. US (United States), 021 (Northern America)
- 61 variants
- 67 redundant
- 26 grandfathered

CUSTOM LANGUAGE TAGS

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

- Private language, e.g.
  - **x-byzantin-Latn**: Byzantine Greek (transliterated)
  - **x-khasian**: Khasian
  - **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-142**: Asian language (not specified which)
  - **qqq-ET**: Ethiopian (not specified which: Boro/Borna, Karo, Male…)
- 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**: Chinese (transliterated Pinyin without tones)

Future: publish lang tags (we now publish only ISO2 & ISO3 codes)

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... 2002-. ".
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; relations, place types (using rdf:Statement)

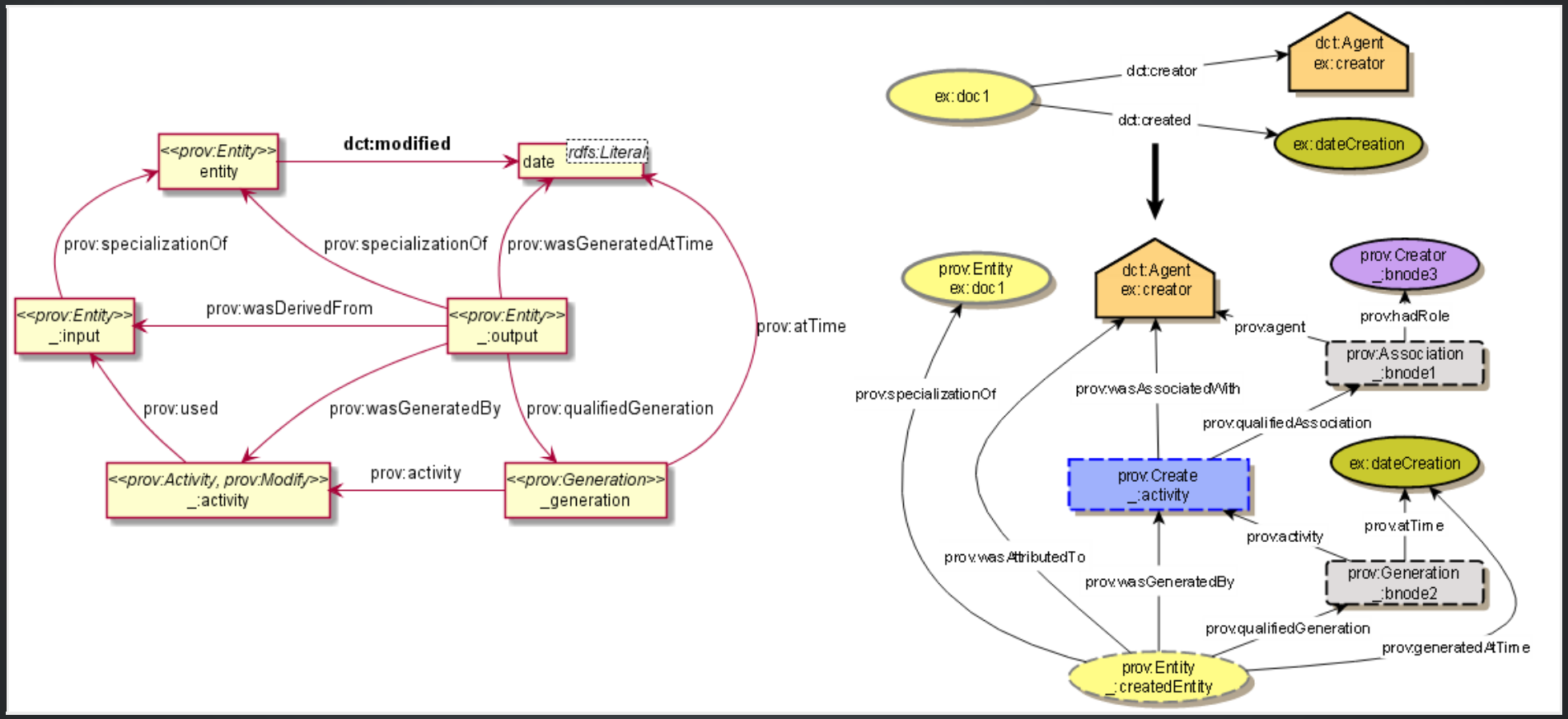
```
aat_term:1000002693-en a skosxl:Label;
  skosxl:literalForm "lambruscatura"@en ;
  gvp:historicFlag <http://vocab.getty.edu/historic/historic> ;
  schema:startDate "0900"^^xsd:gYear ;
  schema:endDate "1700"^^xsd:gYear ;
  rdfs:comment "Medieval term for wainscoting".

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:gYear;
  gvp:displayOrder "1"^^xsd:positiveInteger.
```

PROVENANCE ONTOLOGY

- PROV considers that prov:Modify uses an unknown old entity "_:input" and generates an unknown new entity "_:output", both being specializations of the entity under consideration.
- Need to use prov:Generation so we can use prov:atTime and reflect that the modification is a prov:InstantaneousEvent.



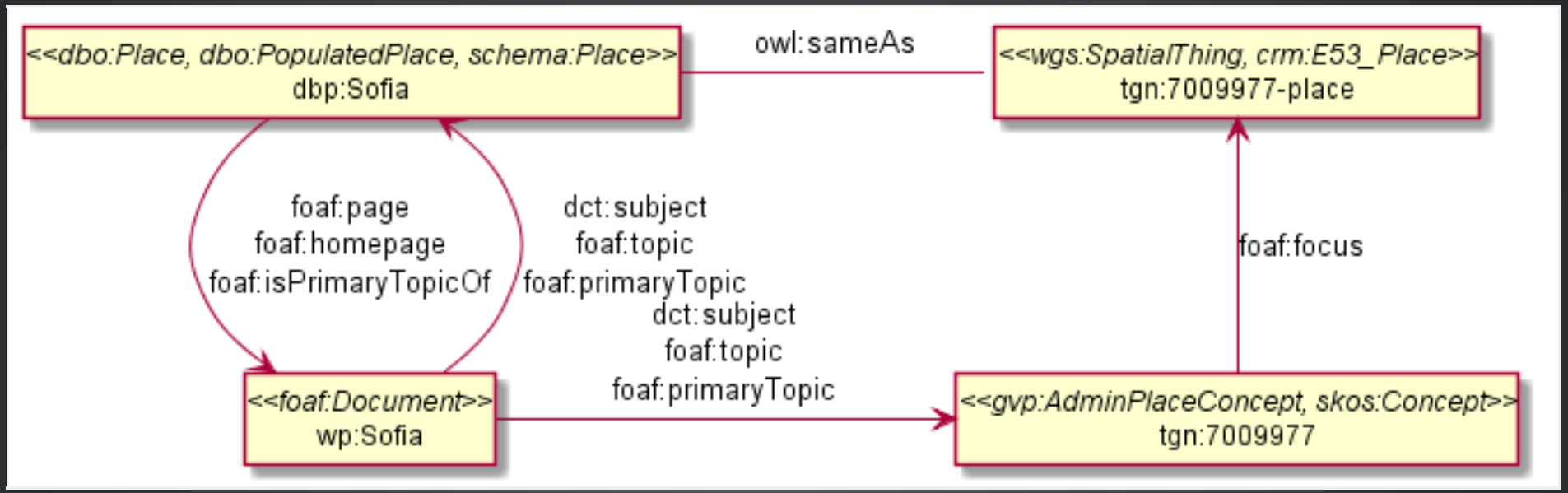
REVISION HISTORY

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.
```

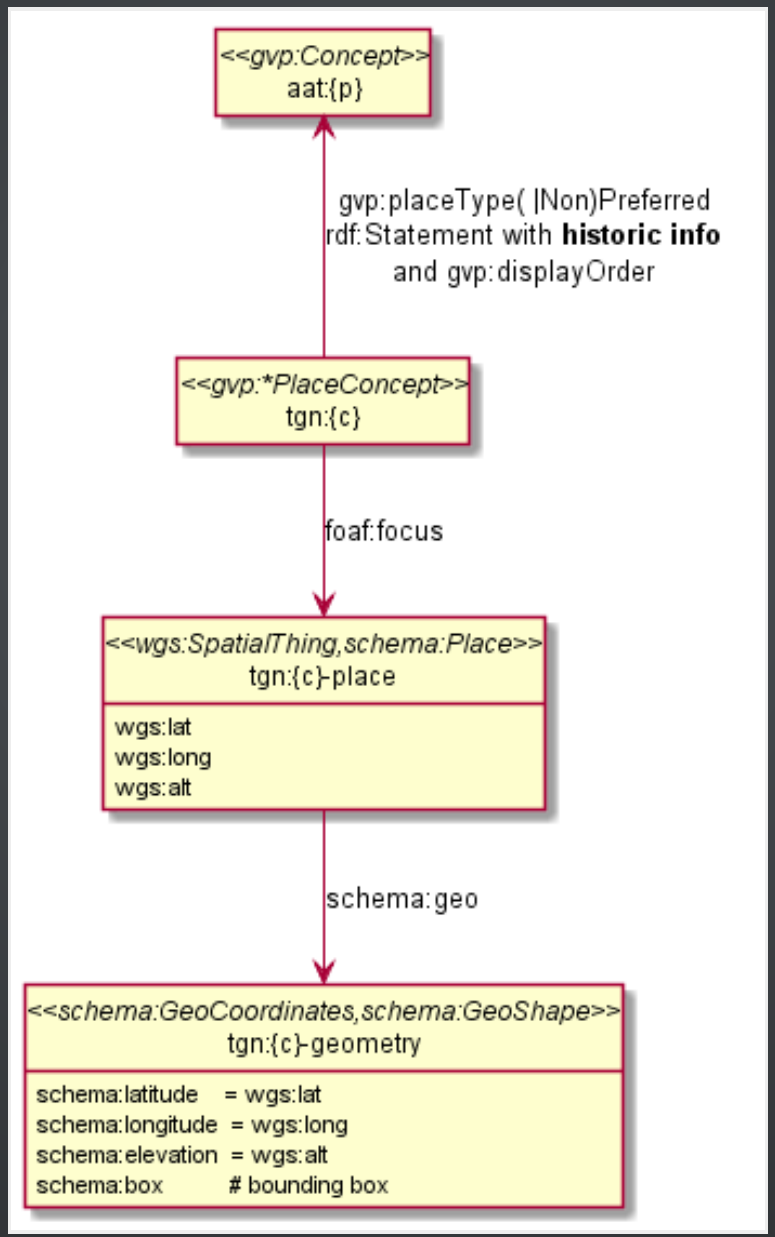
TGN SPECIFICS: CONCEPT-PLACE DUALITY

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



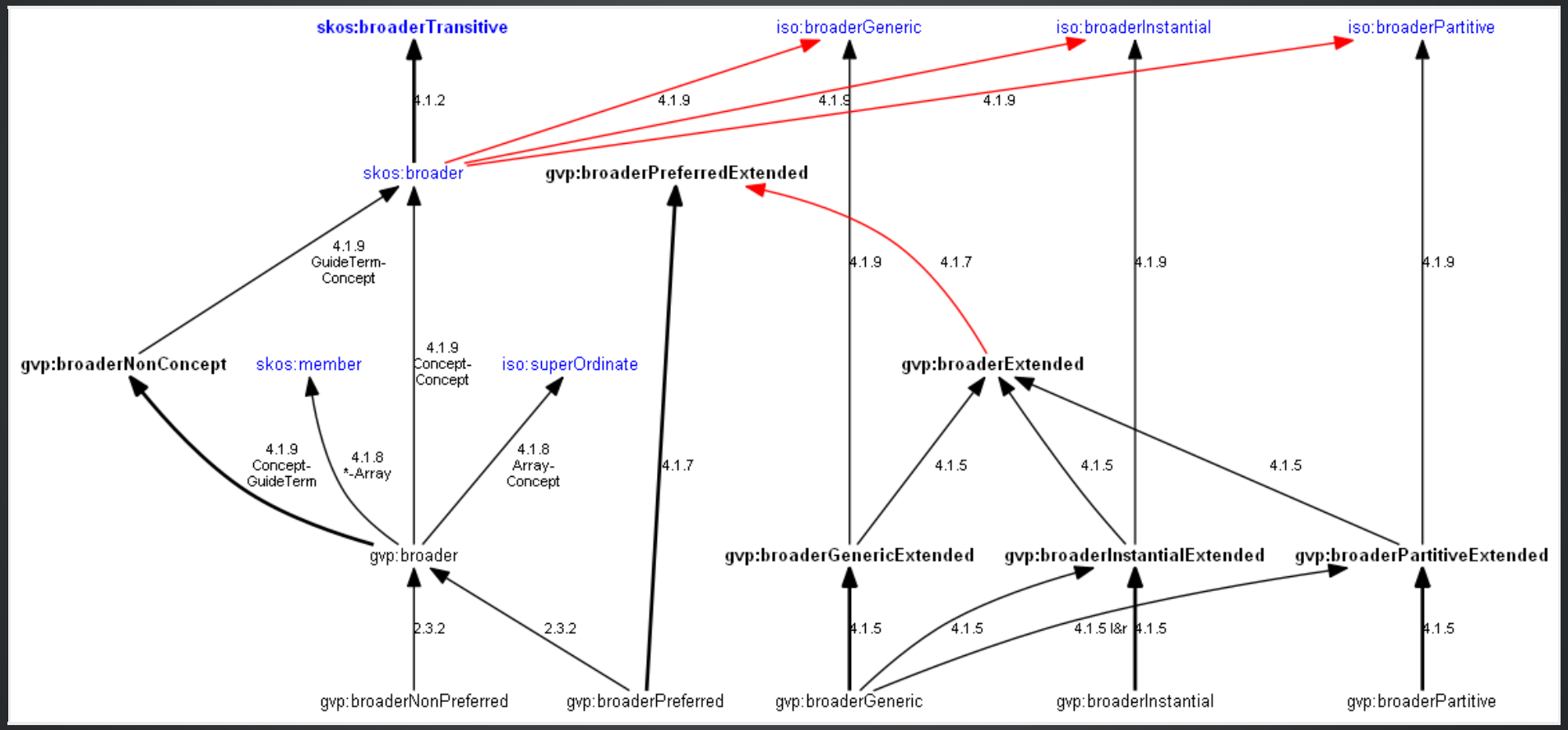
TGN SEMANTIC REPRESENTATION

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



INFERENCE

Hierarchical Relations inference (GVP->Standard):
blue=standard, black=GVP, bold=closure, red=restriction. Numbers refer to doc sections



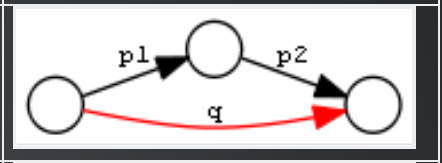
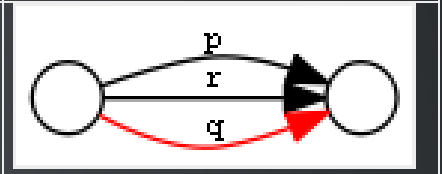
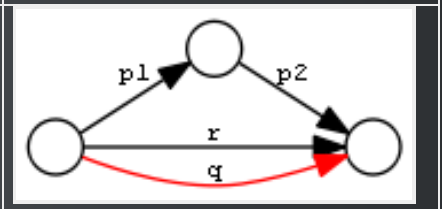
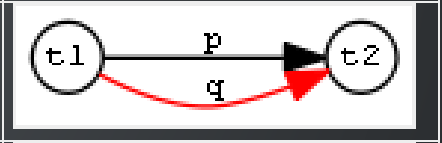
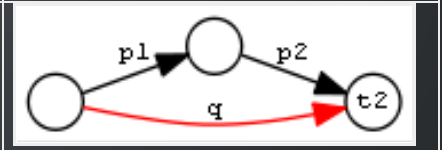
EXTENDED PROPERTY CONSTRUCTS

- TGN is much bigger: AAT: 10M, TGN: 94M (explicit statements)
- We infer 60M statements (1.58x expansion ratio)
- To do this quickly (on biweekly refresh), we decided to use OWLIM Rules
- While OWL2 has very powerful class constructs, its property constructs are quite weak
- Extending OWL2 Property Constructs: several extensions that we found useful

`pN` = premises, `r` = restriction (just another premise), `tN` = types, `q` = conclusion

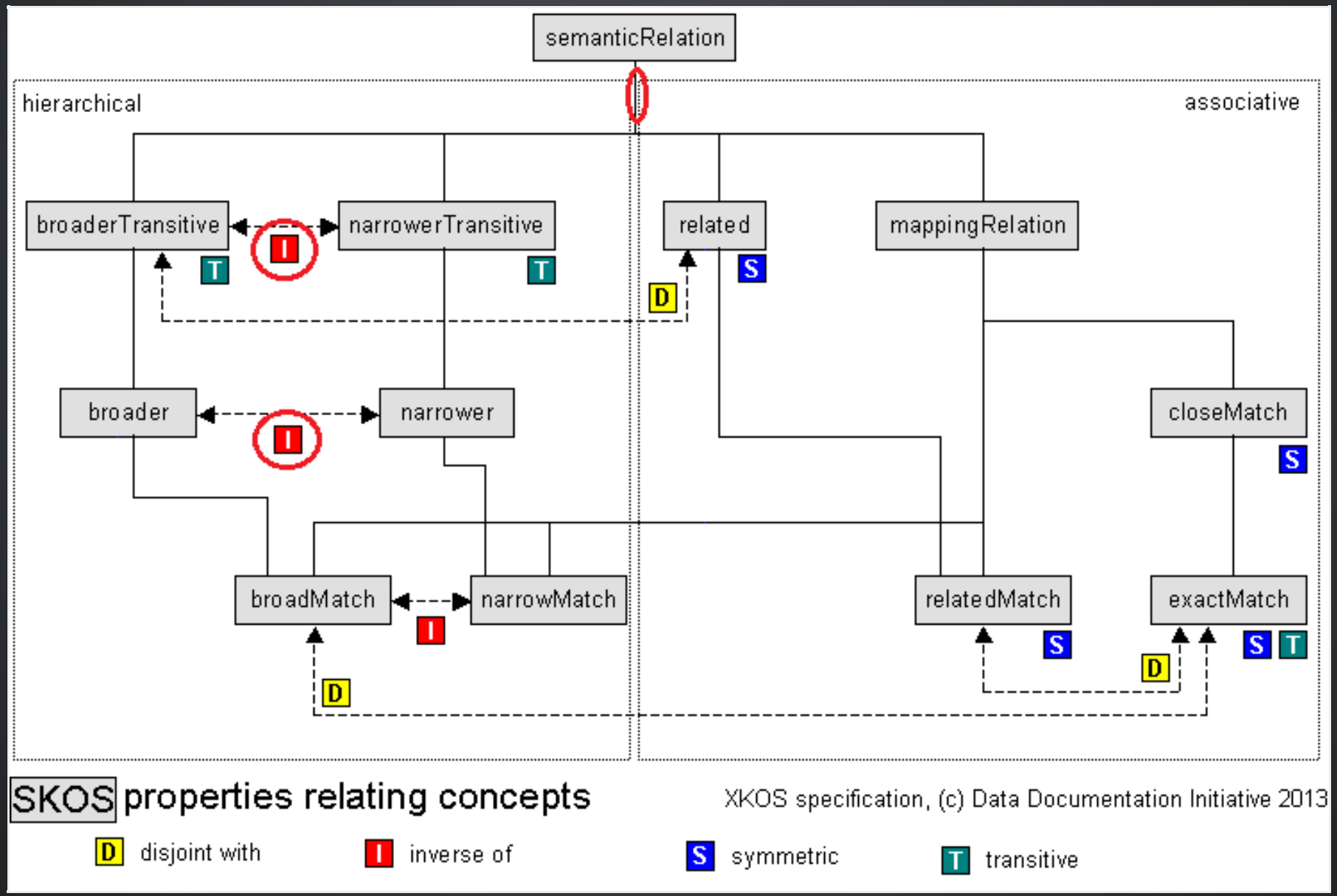
- `p1 / p2`: property chain (more efficient than owl:propertyChainAxiom and owl:TransitiveProperty)
- `p & r`: property conjunction (restriction): holds between two nodes when both properties connect the same nodes
- `[t1] p [t2]`: type restriction: holds when source has type t1 and target has type t2 (shown inside the node)

EXTENDED PROPERTY CONSTRUCTS (2)

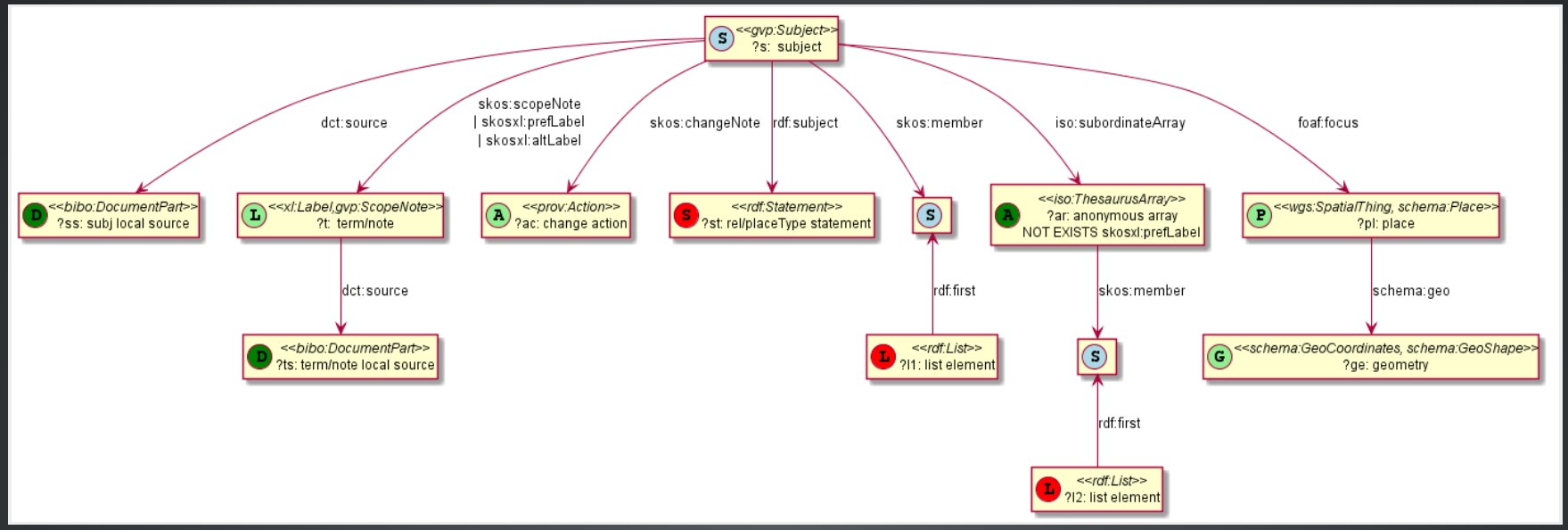
name	prop path	construct	illustration
PropChain	q <= p1 / p2	Chain of fixed length 2	
PropRestr	q <= p & r	Conjunction (restriction by property)	
PropChainRestr	q <= (p1 / p2) & r	Chain and restriction by property	
TypeRestr	q <= [t1] p [t2]	Restriction by two typechecks	
PropChainType2	q <= p1 / p2[t2]	Chain and typecheck	

REDUCED SKOS INFERENCE

Eliminate redundant props (World has 1.2M narrowerTransitive, 2.4M semanticRelation). Break inference at red ovals



CONSTRUCT QUERY: GET & CACHE ALL DATA FOR SUBJECT




```
CONSTRUCT {
  ?s  ?p1 ?o1. # subject
  ?ac ?p2 ?o2. # change action
  ?t  ?p3 ?o3. # term/note
  ?ss ?p4 ?o4. # subject local source
  ?ts ?p6 ?o6. # term/note local source
  ?st ?p7 ?o7. # statement about relations/placeTypes
  ?ar ?p8 ?o8. # anonymous array of subject
  ?l1 ?p9 ?o9. # list element of subject
  ?l2 ?pA ?oA. # list element of anonymous array
  ?pl ?pB ?oB. # place
  ?ge ?pC ?oC. # geometry
} WHERE {
  BIND (tgn:3000034 as ?s)
  {?s ?p1 ?o1 FILTER(!isBlank(?o1))}
  UNION {?s skos:changeNote ?ac. ?ac ?p2 ?o2}
  UNION {?s dct:source ?ss. ?ss a bibo:DocumentPart. ?ss ?p4 ?o4}
  UNION {?s skos:scopeNote|skosxl:prefLabel|skosxl:altLabel ?t.
     {?t ?p3 ?o3 FILTER(!isBlank(?o3))}
     UNION {?t dct:source ?ts. ?ts a bibo:DocumentPart. ?ts ?p6 ?o6}}
  UNION {?st rdf:subject ?s. ?st ?p7 ?o7}
  UNION {?s skos:member/^rdf:first ?l1. ?l1 ?p9 ?o9}
  UNION {?s iso:subordinateArray ?ar FILTER NOT EXISTS {?ar skosxl:prefLabel ?t1}.
     {?ar ?p8 ?o8 FILTER(!isBlank(?o8))}
     UNION {?ar skos:member/^rdf:first ?l2. ?l2 ?pA ?oA}}
  UNION {?s foaf:focus ?pl.
     {?pl ?pB ?oB}
     UNION {?pl schema:geo ?ge. ?ge ?pC ?oC}}
```

SEMANTIC RESOURCES, DUMPS

- **All** data for every independent resource (Subject, Source, Contributor) is extracted
- Using CONSTRUCT queries like above (the ones for Source, Contributor are much simpler)
- Entity files are cached, thus served very quickly
- Entity files are served in RDF/XML, N3/Turtle, NTriples, JSON, soon JSON-LD
- explicit.zip: R2RML-generated statements, NTriples (**you** need to do the Inferencing)
- full.zip: all statements, concatenated from entity files, NTriples

DOCUMENTATION

Getty Vocabularies: Linked Open Data

Very detailed: 100 pages! Linkable anchors:

vocab.getty.edu/doc/#Full_Text_Search

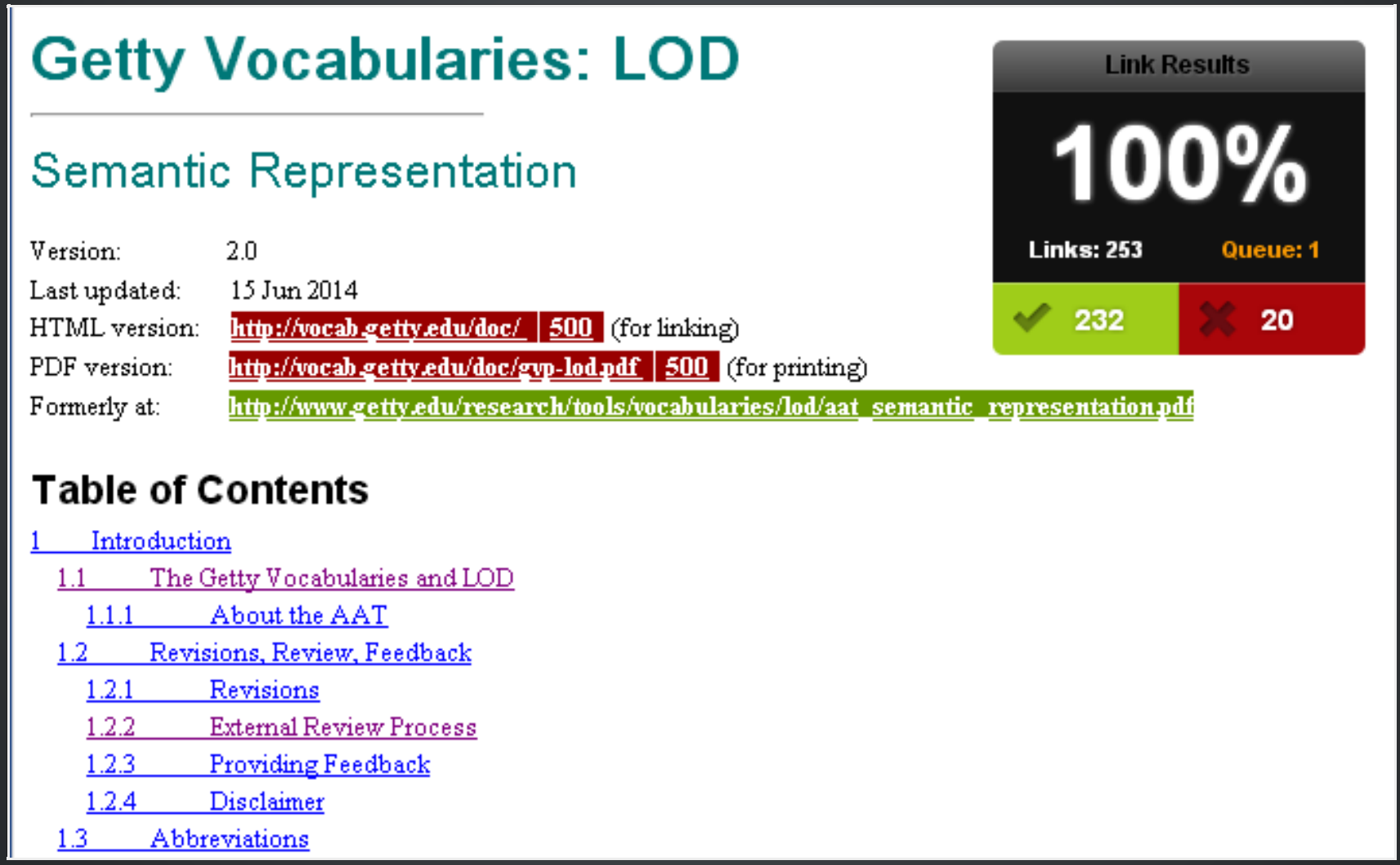
Semantic Representation

Version: 2.0
Last updated: 19 Aug 2014
HTML version: http://vocab.getty.edu/doc/ (for link
PDF version: http://vocab.getty.edu/doc/gvp-lod.
Formerly at: http://www.getty.edu/research/tool
Initial version: Vladimir Alexiev, Joan Cobb, Greg
Updates: Vladimir Alexiev, Joan Cobb

Table of Contents

DOC PRODUCTION

Edit in Word. Spellcheck, link check (below). Print as PDF. Save as Compact HTML, HtmlTidy, rewrite with original images.



Getty Vocabularies: LOD

Semantic Representation

Version: 2.0
Last updated: 15 Jun 2014
HTML version: http://vocab.getty.edu/doc/ 500 (for linking)
PDF version: http://vocab.getty.edu/doc/gvp-lod.pdf 500 (for printing)
Formerly at: http://www.getty.edu/research/tools/vocabularies/lod/aat_semantic_representation.pdf

Link Results
100%
Links: 253 Queue: 1
232 20

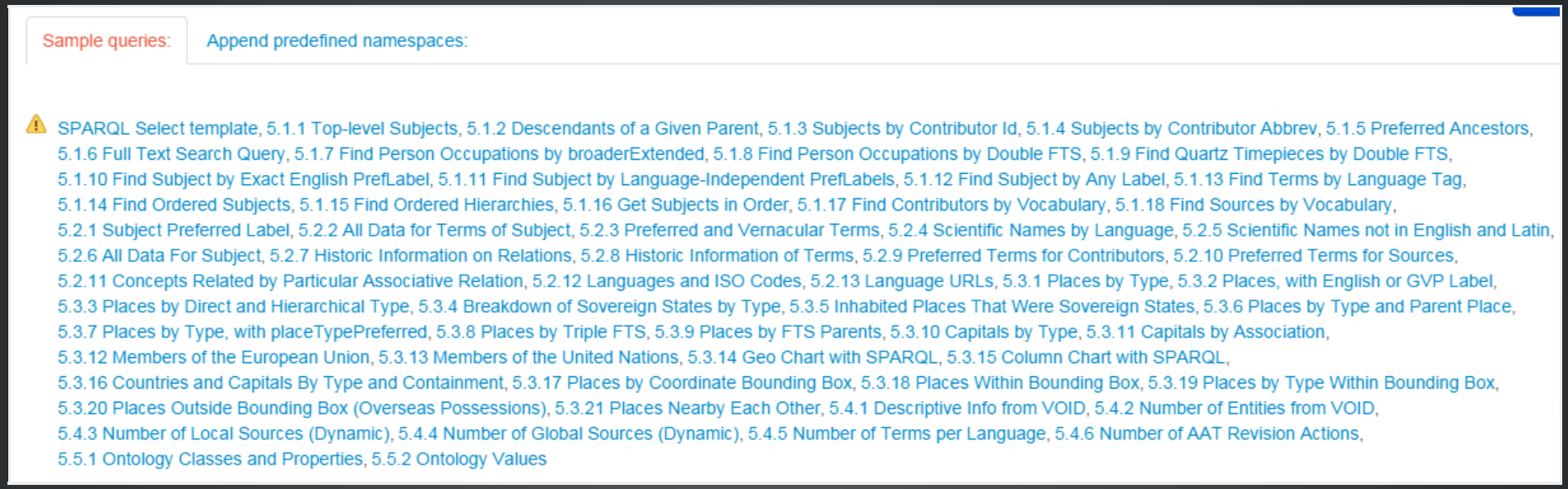
Table of Contents

1 Introduction
1.1 The Getty Vocabularies and LOD
1.1.1 About the AAT
1.2 Revisions, Review, Feedback
1.2.1 Revisions
1.2.2 External Review Process
1.2.3 Providing Feedback
1.2.4 Disclaimer
1.3 Abbreviations

Benefit: printable PDF and **linkable** HTML

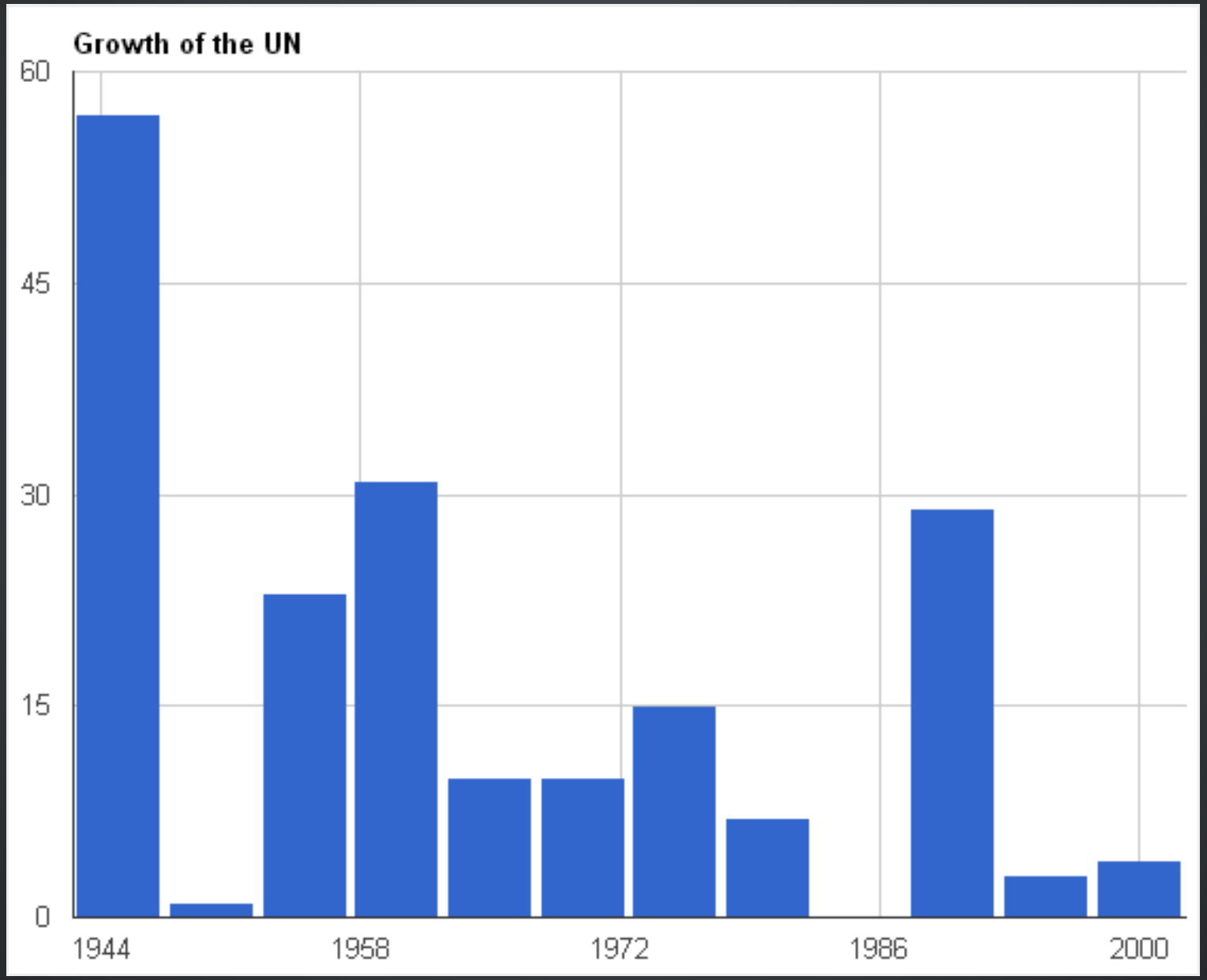
SAMPLE QUERIES

Lots of them! The (!) says "read the documentation first". As part of helpdesk support, we're tracking usage and adding samples.



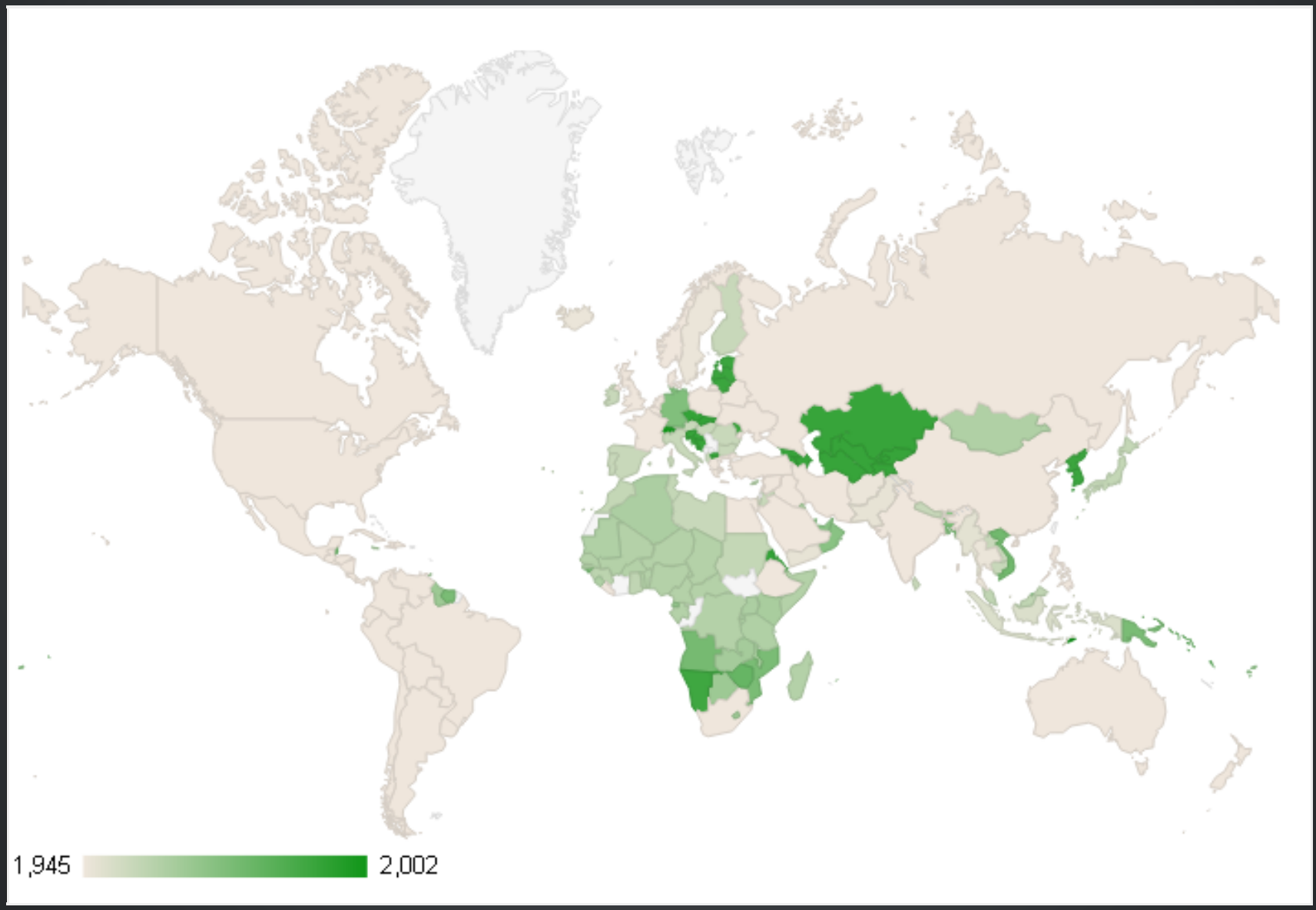
SAMPLE QUERY: BAR CHART WITH SPARQL

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



SAMPLE QUERY: GEO CHART WITH SPARQL

When each nation joined the UN. See doc or jsfiddle with it.



SAMPLE QUERY: OVERSEAS POSSESSIONS OF THE NETHERLANDS

```
# 5.3.20 Places Outside Bounding Box (Overseas Possessions)
select ?place ?name ?lat ?long {
  ?place skos:inScheme tgn: ;
    foaf:focus [wgs:lat ?lat; wgs:long ?long];
    gvp:prefLabelGVP [xl:literalForm ?name];
    gvp:broaderPartitiveExtended [rdfs:label "The Netherlands"@en]
  filter (!(50.787185 <= ?lat && ?lat <= 53.542265 && 3.389722 <= ?long && ?long <= 7.169019))}
```

Results for # 5.3.20 Places... (100 of 596)

Download

place	name	lat	long
tgn:7256571	Back Off Bay@nl	17.45	-62.95
tgn:7005674	Philipsburg	18.05	-63.0833
tgn:1011622	Basora, Punta	12.4167	-69.85

GVP LOD USAGE

People started using AAT and TGN right after their release

- AAT Concept selection (usually by autocompletion): EADitor, xEAC, VRA Editor, MODES, DIGIMUS, Drupal Web Taxonomy
- TGN Place selection: same as above; Portable Antiquities (finds.org.uk), Nomisma, Kerameikos
- Visualization: Hierarchies with d3js, LOD with lodlive.it
- Semantic enrichment: Europeana (Rijksmuseum, Museo Galileo, Erfgoedplus.be), Partage Plus (Art Nouveau to Europeana)

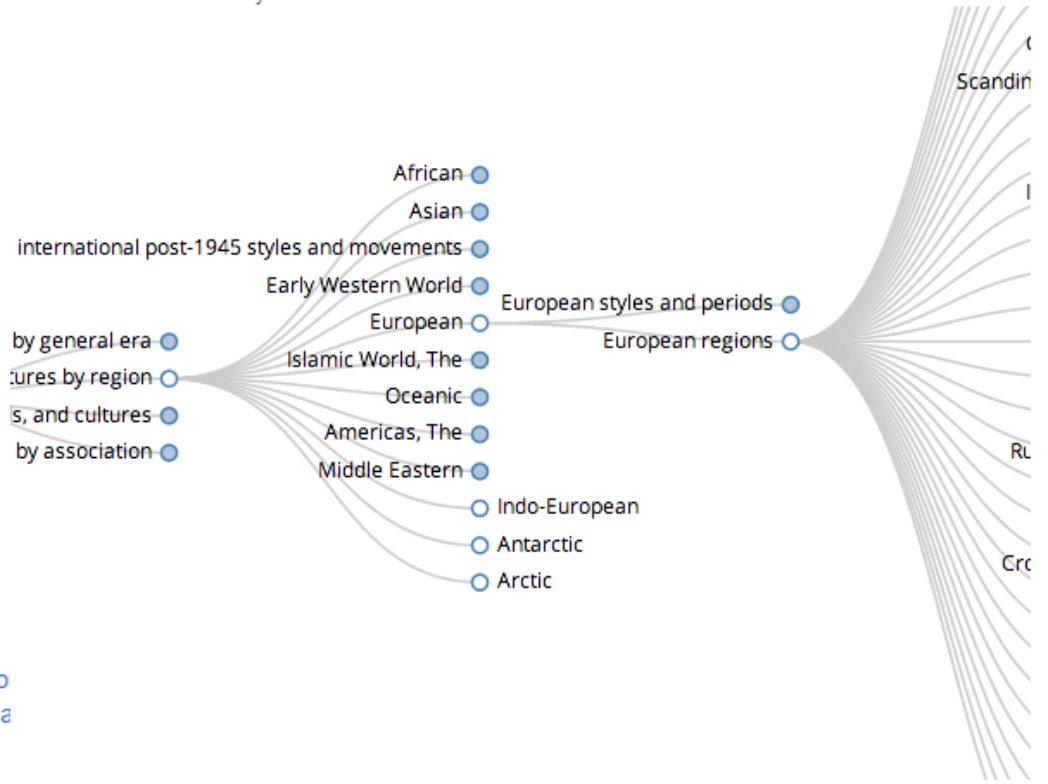
Vocab	**Actual**	**Potential**
AAT	9	6
TGN	3	
ULAN		1

USAGE STORIES

Usage Stories (internal confluence)

Short descriptions of ways to make use of AAT once it is launched as LOD. Or even before it's officially launched!

- AAT Actual Usage
  - Use 1 Integrating EADitor with Getty linked data AAT
  - Use 2 Using AAT in VRA XForms editor
  - Use 3 Visualizing Hierarchies with d3js
  - Use 4 Visualizing with en.lodlive.it
  - Use 5 AAT as a MODES web termlist
  - Use 6 AAT as Categories in Wikidata Visual Arts project
  - Use 7 AAT Classification in DIGIMUS
    - DIGIMUS object (Astrolabe)
    - DIGIMUS Material
    - DIGIMUS Period
    - Comments to DIGIMUS
  - Use 7 Drupal Web Taxonomy plugin for Getty vocabularies
  - Use 8 AAT in Europeana
  - Use 9 AAT in Partage Plus
- AAT Potential Usage
  - Story 1 Using GVP LOD in embedded photo metadata
  - Story 2 LOD for sharing public information
  - Story 3 Joining collections
  - Story 4 Current use of controlled vocabularis enabled future explo
  - Story 5 Enriching an entire culture data ecosystem like Europeana
  - Story 6 Using GVP LOD for Digital Art History
- TGN Actual Usage
  - Use 21 TGN in the Portable Antiquities Scheme
  - Use 22 TGN in Nomisma and Kerameikos
  - Use 23 TGN as GeoJSON
- ULAN Potential Usage
  - Story 41 Analyze ULAN Networks



THANKS FOR YOUR TIME!

If you have any questions or suggestions for improvement, please don't hesitate to contact me: vladimir.alexiev@ontotext.com