

Gramm Richardson - U.S. Department of Defense
Elli Schwarz - SRA International, Inc.

30 May, 2012

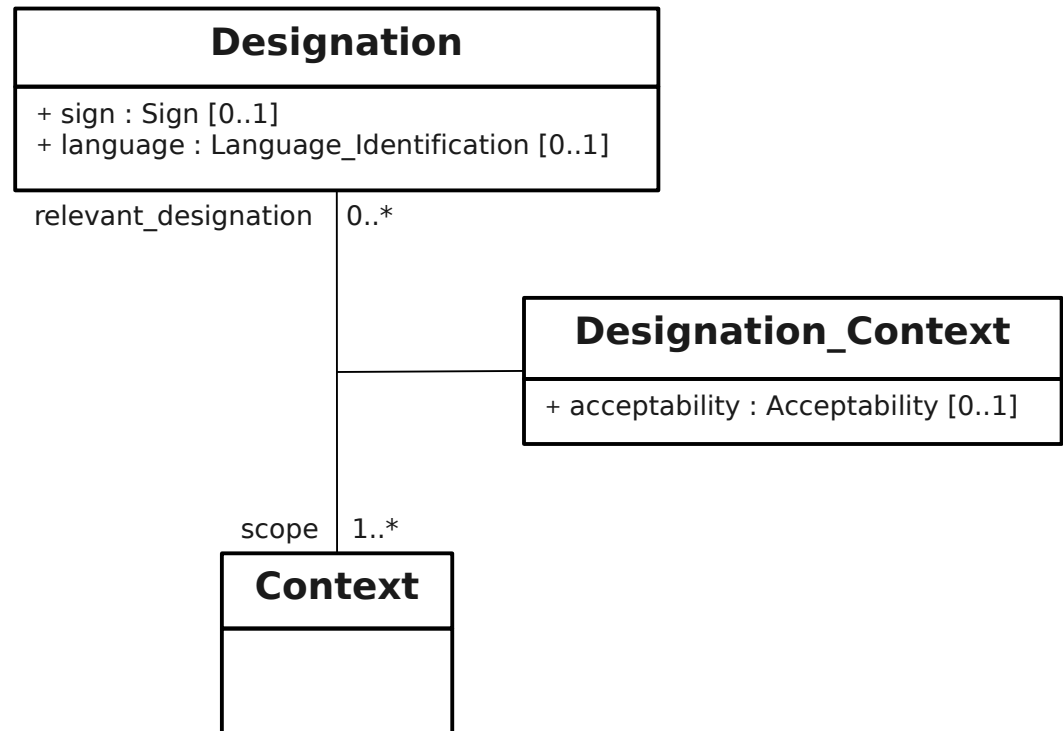
ISO/IEC 11179-3 Metamodel OWL Ontology

- 58 Classes
- 74 Datatype properties
- 113 Object properties
- Inspired by Kevin Keck's 11179 OWL Ontology
- Implementation Decision
 - XML Schema Datatypes
 - Implementation-specific subclassing
- Implemented
 - Concepts metamodel region
 - Conceptual and Value_Domain region
 - Designation and Definition metamodel region

Assertion
Link
Classifiable_Item
Classification
Concept
 Conceptual_Domain
 Described_Conceptual_Domain
 Enumerated_Conceptual_Domain
 Data_Element_Concept
 Dimensionality
 Object_Class
 Property
 Relation
 Binary_Relation
 Relation_Role
 Unit_of_Measure
 Value_Meaning
Concept_System
Contact
 Registrar
Context
Data_Element
Data_Element_Derivation
Data_Element_Example
Datatype
Definition
Definition_Context
Derivation_Rule
Designatable_Item
Designation
Designation_Context
Document_Type
Identified_Item
 Registered_Item
 Administered_Item
 Attached_Item
Individual
Language_Identification
Link_End
Namespace
Naming_Convention
Organization
 Registration_Authority
Permissible_Value
Reference
Reference_Document
Registration
Registration_Authority_identifier
Registration_State
Registry_Specification
Role
Scoped_Identifier
Slot
Stewardship_Record
Submission_Record
Value_Domain
 Described_Value_Domain
 Enumerated_Value_Domain

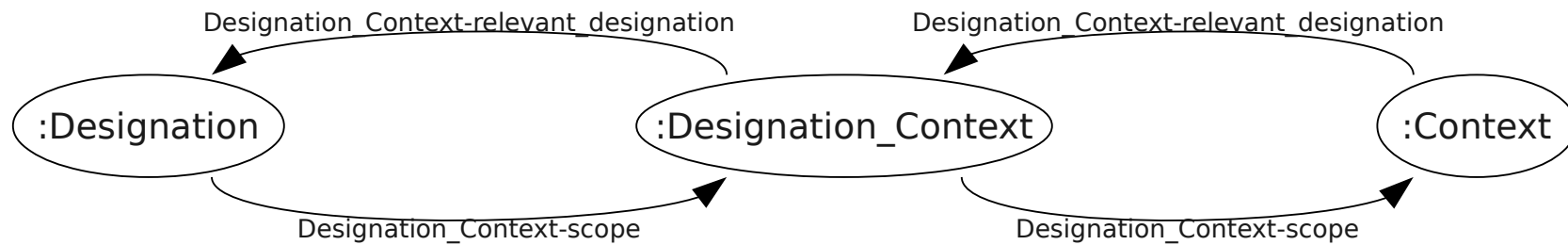
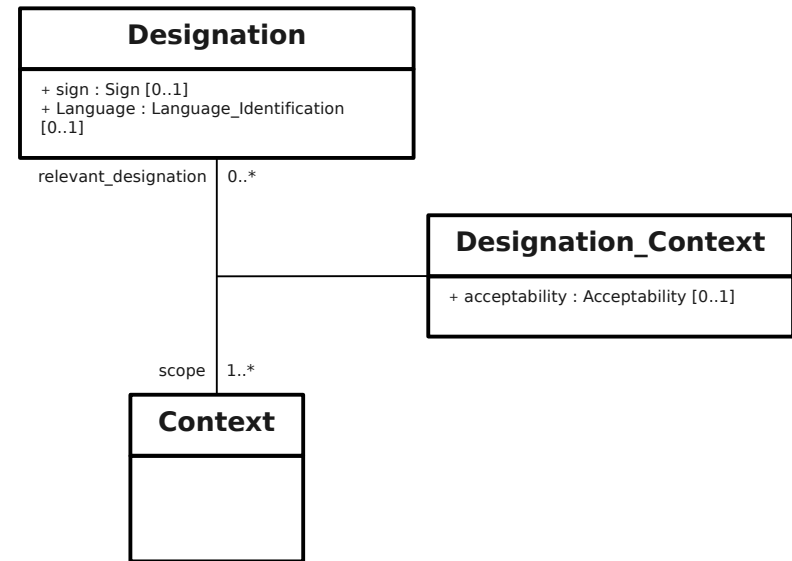
ISO/IEC 11179-3 Metamodel OWL Ontology

- Association classes
- Implemented in OWL 1 and RDF as a class with additional properties



ISO/IEC 11179-3 Metamodel OWL Ontology

- Association classes
- Implemented in OWL 1 and RDF as a class with additional properties
- OWL object properties created from a combination of metamodel association and role



ISO/IEC 11179-3 Metamodel OWL Ontology

- begin_date and end_date extension for Designation_Context
- at first, an accommodation for difficulties in removing RDF triples from the graph store
- Supports management of the history of a term's acceptability in a context

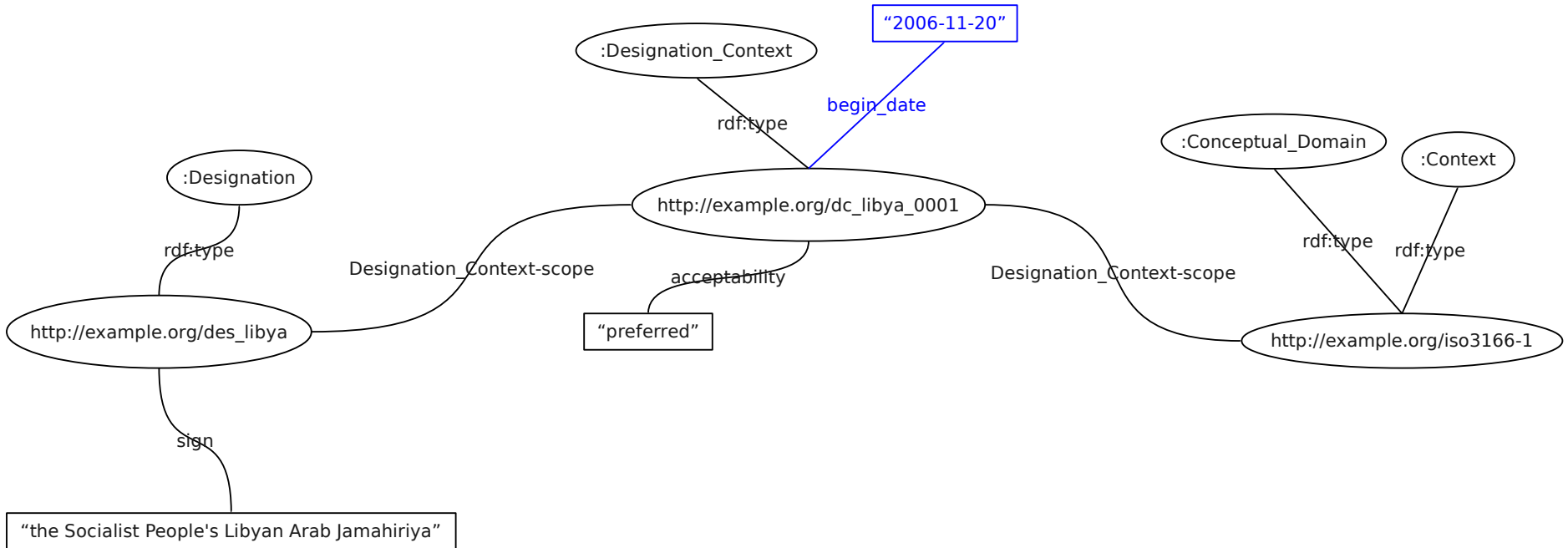
Designation_Context

- + acceptability : Acceptability [0..1]
- + begin_date : Date [1]
- + end_date : Date [0..1]

ISO/IEC 11179-3 Metamodel OWL Ontology

- begin_date and end_date extension for Designation_Context

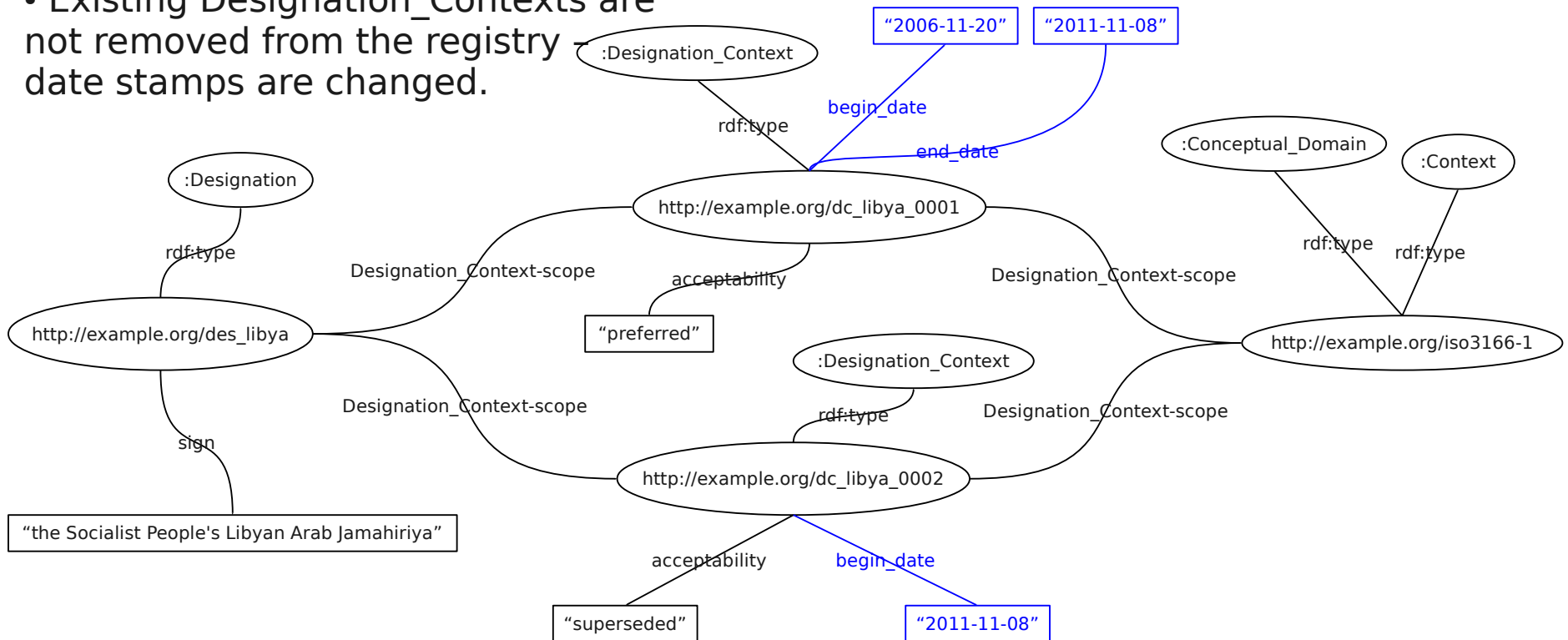
Designation_Context
+ acceptability : Acceptability [0..1]
+ begin_date : Date [1]
+ end_date : Date [0..1]



ISO/IEC 11179-3 Metamodel OWL Ontology

- Updating the Designation_Context is a strictly additive process.
- Conceptual_Domain as the Context for Designations
- Existing Designation_Contexts are not removed from the registry date stamps are changed.

Designation_Context
+ acceptability : Acceptability [0..1]
+ begin_date : Date [1]
+ end_date : Date [0..1]



Country Code Sets as Conceptual_Domains and Value_Domains

- Need a methodology that can be generalized to multiple code sets.
- Concept_Systems for first order or second order geopolitical entities
- Working with ≈ 12 code sets, designations in 8 languages fully covering one or more code sets

ISO 3166-1, 2

**National Olympic
Committees**

UIC 920-14

**GSA
Locator Codes**

ITU-T e.212

**Census
Schedule C**

ITU-T e.164

**NGA
Geopolitical Codes**

FAOSTAT

**ICAO
Nationality Marks**

AGROVOC

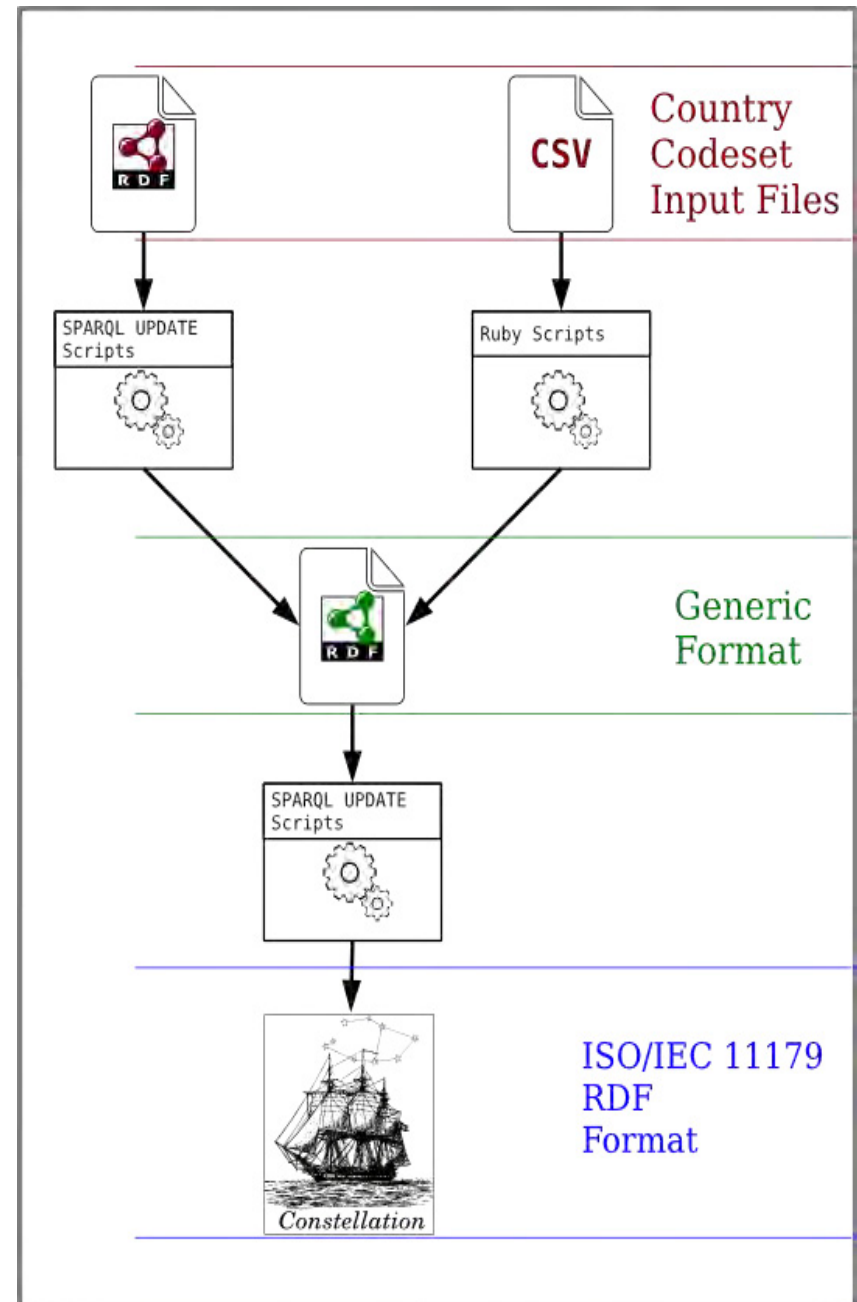
UN M.49

GAUL

**Treasury
International Capital**

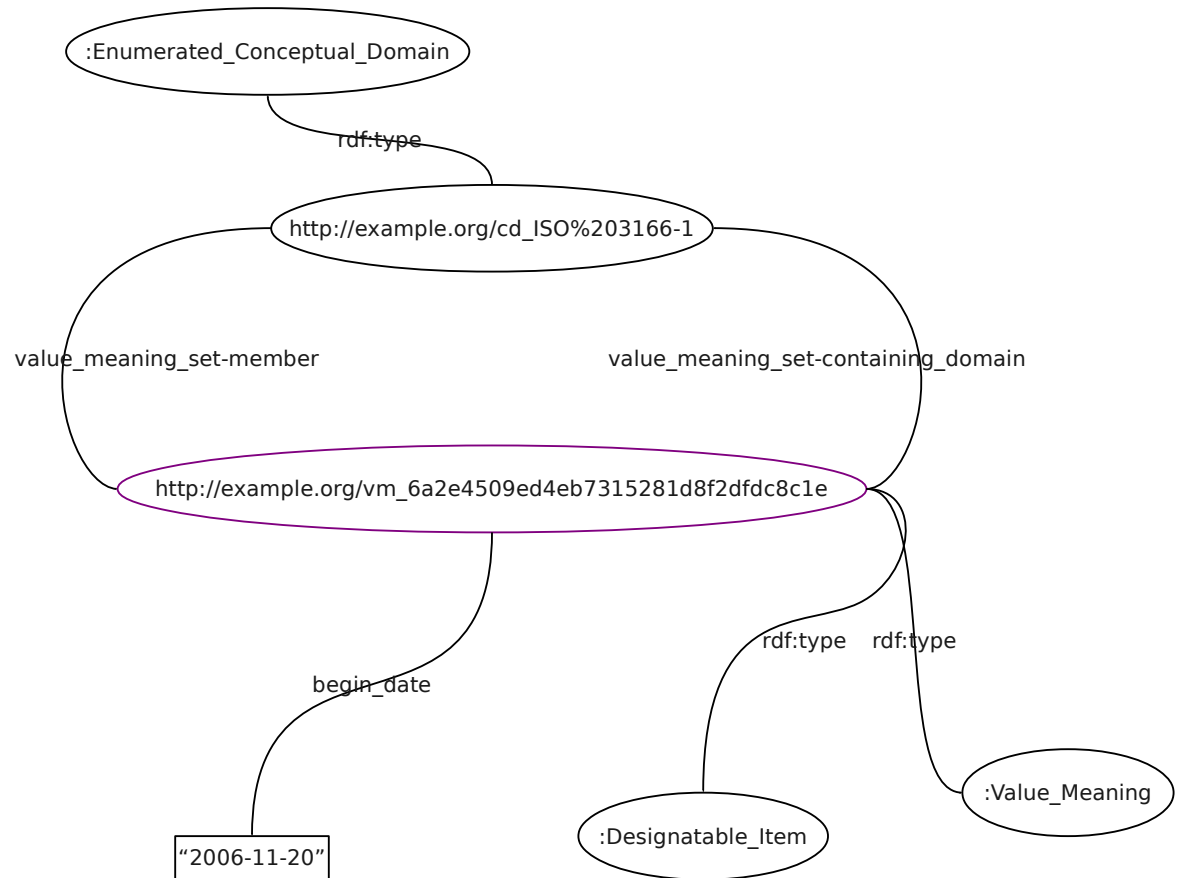
Implementation

- Able to ingest any code sets in RDF format or CSV format with established columns
- Intermediate/generic RDF format
 - Makes it easier to have one workflow to transform data to 11179 format
 - Puts the data in a uniform state for easy comparison
- Store processed code sets in 11179-formatted metadata in RDF Triple Store



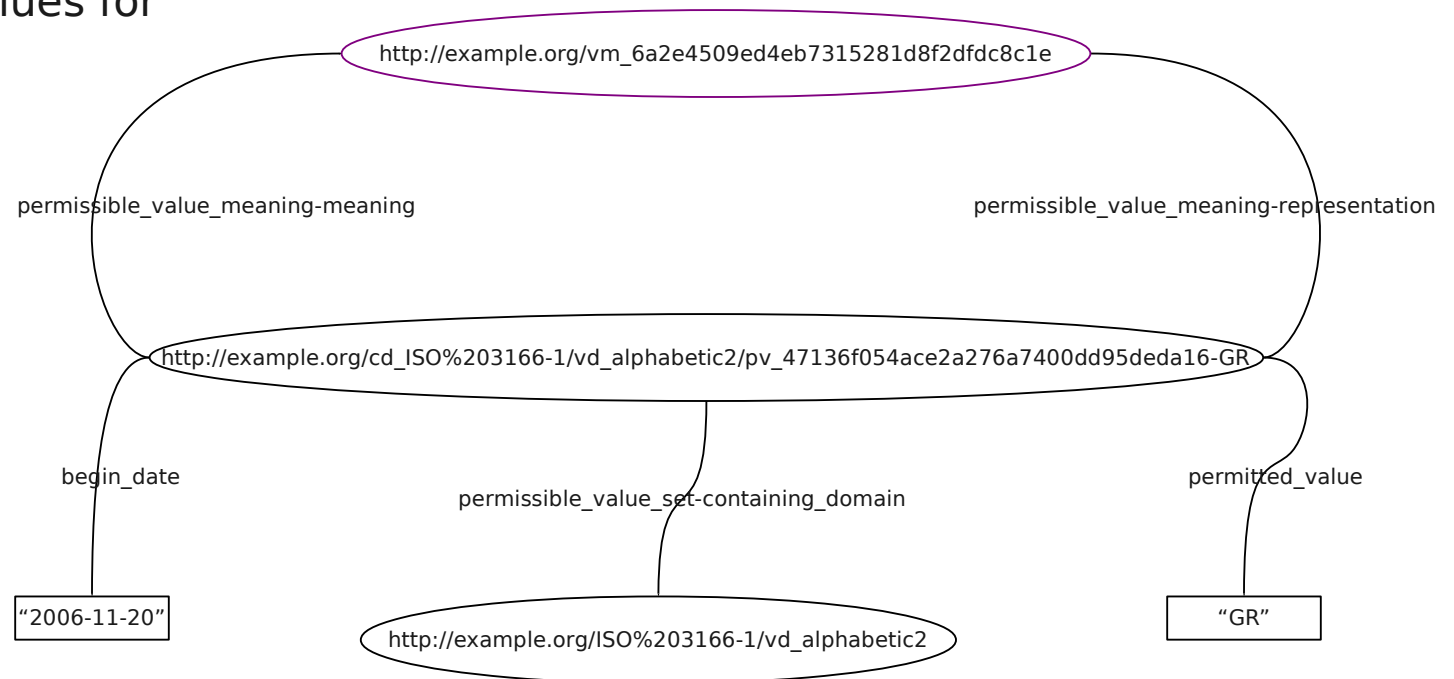
ISO 3166-1:2006 Implementation

- One Conceptual_Domain,
 - alphabetic2
 - alphabetic3
 - numeric3
- Value_Meanings are also typed as Designatable_Items



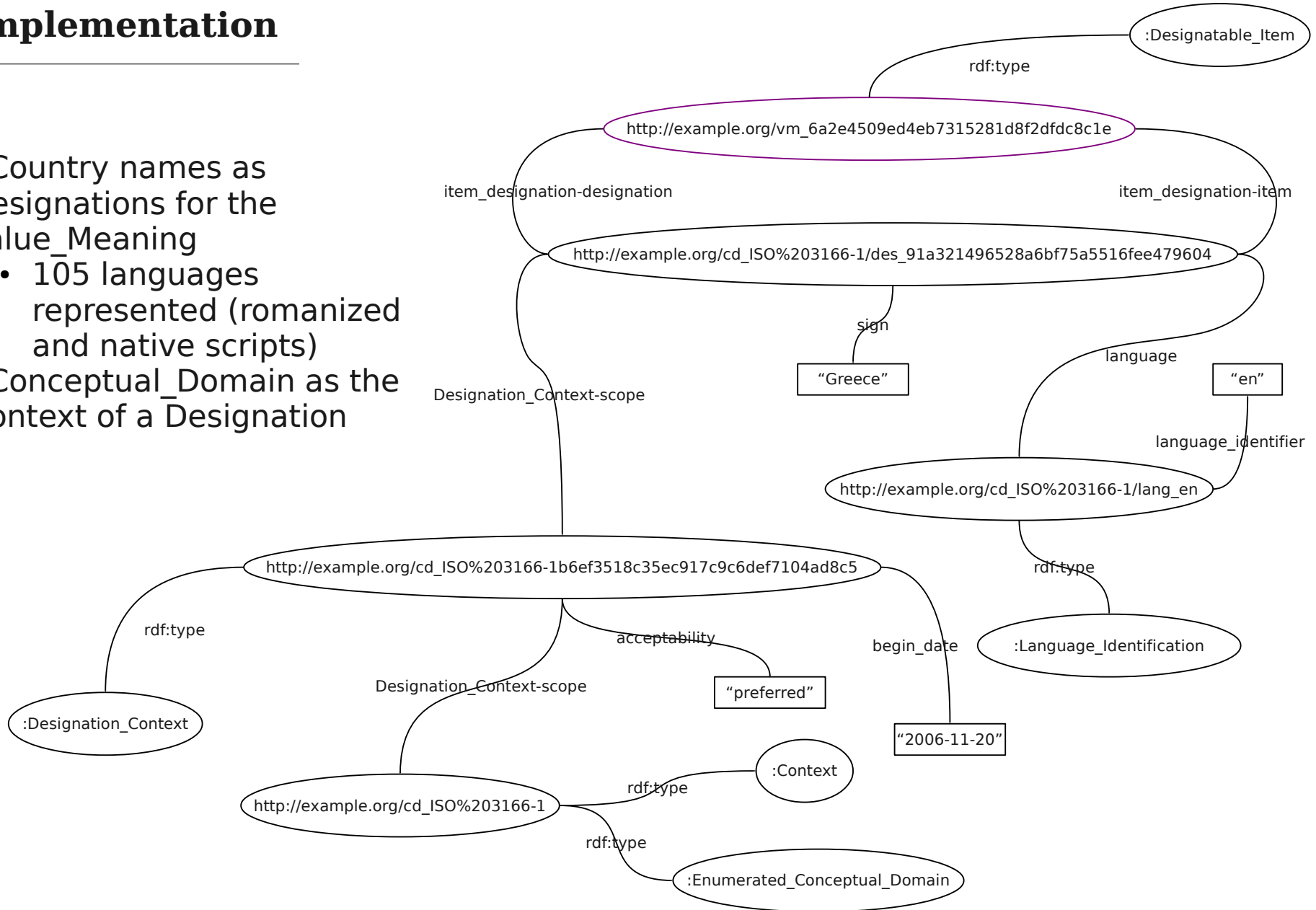
ISO 3166-1:2006 Implementation

- URIs for Permissible_Values are constructed using the path for the Value_Domain URI
- Decided to represent names of countries as Designations rather than Permissible_Values
- Used Permissible_Values for the country codes.



ISO 3166-1:2006 Implementation

- Country names as Designations for the Value Meaning
 - 105 languages represented (romanized and native scripts)
- Conceptual_Domain as the Context of a Designation



Metamodel Implementation Challenges

- Indirection and complex 11179 metamodel creates problems for
 - Query performance
 - Query maintainability
- For example
 - Relations and Links
 - Designatable_Items
 - Multi-page SPARQL queries to account for optional class associations and attributes.
- SKOS reimplementations as a 11179 Concept_System
 - Link-centric vs. Role-centric
 - Representation of Links and Relations has changed as ISO/IEC 11179-3 FDIS sneak peeks were released

```
cmdr:item_designation-item ?country,  
rdf:type cmdr:Designation.  
?country a cmdr:Value_Meaning;  
rdf:value ?vmValue;  
cmdr:permissible_value_meaning-represe  
cmdr:value_meaning_set-containing_dom  
cmdr:begin_date ?countryBeginDate .
```

```
OPTIONAL {  
  ?country cmdr:concept_system_membership-includi  
  a <http://constellation.jhuapl.edu/countryM  
  rdf:value ?matchUUID  
}  
}
```

```
OPTIONAL {  
  ?country2 cmdr:language [cmdr:language_identifier  
}  
}
```

```
filter(!BOUND(?lang2) || ?lang2 = "en")
```

```
BIND ( concat("results-frame.html?  
queryName=DesignationQuery&showHeader=true&$  
m=", ENCODE_FOR_URI(str(?country2)) ) as ?country2  
}
```

```
OPTIONAL { ?country cmdr:end_date ?countryEnd  
cmdr:begin_date ?countryBeginDate  
}
```

```
#In order for hyperlinks to work with groups, we need  
as a URL and not convert it using the linkformatter in  
BIND ( concat("<a href=", "results-frame.html?  
queryName=GenericQuery&showHeader=true&$Data  
yURI=", ENCODE_FOR_URI(str(?faoURL)), "'>",?vmValu
```

```
BIND ( if(bound(?hyperLink), ?hyperLink, ?country) as
```

```
BIND ( concat("results-frame.html?  
queryName=DesignationQuery&showHeader=true&$  
m=", ENCODE_FOR_URI(str(?country)) ) as ?countryLin
```

```
BIND ( concat("results-frame.html?
```

Benefits of ISO/IEC 11179 Metamodel

- Representing countries using Value_Meanings belonging to Enumerated_Conceptual_Domains, and having Designations and Permissible_Values creates a common framework for promoting data understanding
- Representing code sets in 11179 OWL ontology further facilitates:
 - Uniform data representation
 - Querying across diverse data sources
 - Avoiding terminology conflicts
- Flexibility of the metamodel supports easily adding/modifying information (new code set elements, relationships, updating names or codes for countries)



Postprocessing - Country Matching System

- Semiautomated
- Based on normalized names and trigram similarity with manually-curated truth data
 - “Côte d'Ivoire” \rightleftharpoons “COTE D'IVOIRE”
 - “United States of America” \rightleftharpoons “United States”
 - “Republic of Korea” \rightleftharpoons “South Korea”
 - “Burma” \rightleftharpoons “Myanmar”
- Using skos:closeMatch implemented as a 11179 Relation to associate similar countries
- Each set of matched countries is a Concept_System. The elements in the Concept_System are Value_Meanings that are related to each other via skos:closeMatch.



SPARQL Query System

- SPARQL query templates
 - Provide multiple views of the data
 - Grouping by matched countries
 - Querying by code or name
 - Querying for names in multiple languages or acceptabilities
 - Querying by transnational grouping
 - Specific variables can be parameters for REST services (code, conceptual_domain, country_name)
- Data analysis
- SPARQL 1.1



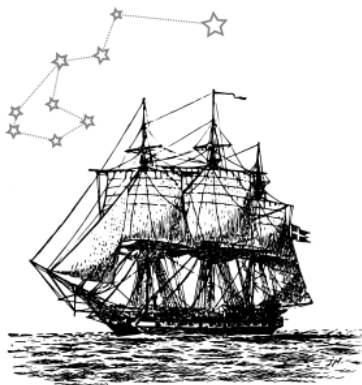
Modeling Challenges

- Difference of opinion between code sets on what geopolitical entities are 'first order' and which are 'second order'
 - U.S. Virgin Islands assigned different codes in ISO 3166-1 and ISO 3166-2
 - British Virgin Islands absent in 3166-2
- Semantics of identical terms can vary across code sets
 - 'unknown', 'undefined', 'reserved'
 - China (whether Hong Kong or Macao are assigned their own codes)



Future Direction

- Implement the Registry metamodel region
 - Research vs. Productization
- Update the ontology to edition 3 IS and use OWL 2 punning.
 - Investigate SPARQL query implementation and maintainability impact.
- Research the management of other types of metadata artifacts.



Constellation