

Naming and Labeling in the Multilingual Web of Data

Daniel Vila-Suero

Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo sn, 28660 Boadilla del Monte, Madrid

<http://www.oeg-upm.net>

dvila@fi.upm.es

Acknowledgements: BabelData Project (TIN2010-17550), Elena Montiel-Ponsoda, Elena Escolano, Boris Villazón-Terrazas, Gordon Dunsire, Asunción Gómez-Pérez, Jorge Gracia

W3C Multilingual Web workshop:
Making the multilingual web work
Rome, 13.03.2013

- Based on

"Style guidelines for naming and labeling ontologies in the multilingual Web" Montiel-Ponsoda, Vila-Suero, Villazón-Terrazas, Dunsire, Escolano and Gómez-Pérez. DC Conference 2011

- + some practical examples/issues from:
 - <http://datos.bne.es>, and
 - IFLA vocabularies translation into Spanish

NAMING

Some general URI design guidelines

Guideline	Source
Don't put too much semantics in the URI	(Berners-Lee, 1998)
Don't bind URIs to some classification or topic (you may change your point of view)	(Berners-Lee, 1998)
Use short URLs as much as possible	(Théreaux, 2003) *
Choose a case policy (all lowercase or first letter uppercase)	(Théreaux, 2003)
Use lower case	(Cabinet Office, 2009)
Separate words by hyphens	(Cabinet Office, 2009)
Where the reference local name may change, or becomes overly verbose, a code may be more appropriate	(Cabinet Office, 2009)

Naming: Preliminary guidelines for a multilingual scenario

Guidelines for the Multilingual Web	Example
Use preferably opaque URI local names	http://aims.fao.org/aos/agrovoc/c_330988
When using meaningful URI local names, use your own language with diacritics, etc.	http://geo.linkeddata.es/ontology/rio
When using meaningful URI local names, use dash or CamelCase for word delimiter, depending on what reads more easily in your own language.	http://geo.linkeddata.es/ontology/aguas-corrientes

Some tools are not prepared for opaque URIs (Pubby)...

* <http://datos.bne.es/resource/XX1718747>

ifla-frbr:P3039	▪ Cervantes Saavedra, Miguel de
ifla-frbr:P3040	▪ 1547-1616
ifla-frad:P4031	▪ Cervantes ▪ Cervantes di Saavedra, Michele ▪ Cervantes, Miguel de ▪ Servantes Saavedra, Migel ▪ Sîrfantis Saafedrâ, Mîgîl dî ▪ Therbantes, Minkel nte ▪ Zerbantes eta Saabedra, Mikel ▪ Θερβάντες, Μιγκέλ ντε ▪ Сервантес Сааведра, Мигель де ▪ سرفانتس ساافدرا، میگیل دی

Semantic Web Journal reviewer about datos.bne.es' paper* :

"It is pity that local names of chosen IFLA-FRBR properties are cryptic codes ... but **authors of this paper are not to blame about that**"

* <http://www.semantic-web-journal.net/content/datosbnees-library-linked-data-dataset>

Some others are better prepared (Puelia)...

* <http://datos.bne.es/frontend/persons>

Display labels are configurable using a Turtle config file

tiene fechas de persona	1547-1616	
tiene nombre de persona	Cervantes Saavedra, Miguel de	
tipo	Person / Persona	

Label not selected based on User's locale

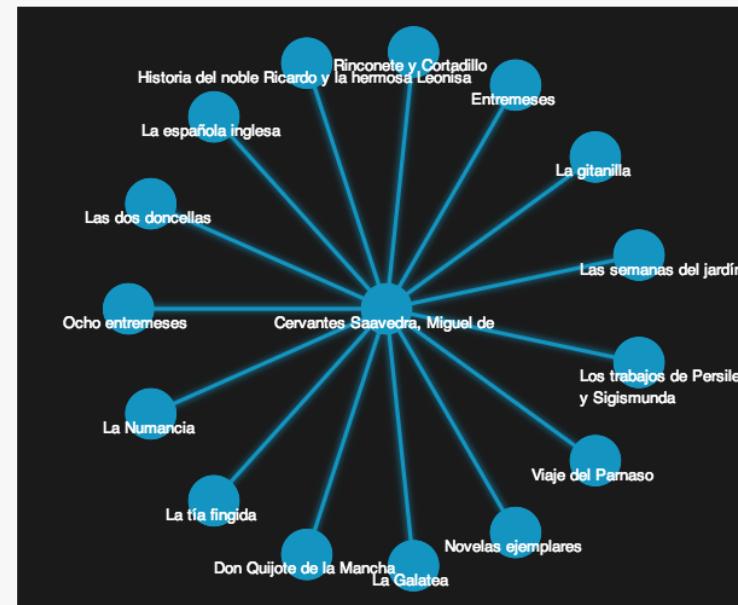
frbr:C1005 a rdfs:Class;
rdfs:label "Person"@en, "Persona"@es

Some personal experiences using opaque URIs

The screenshot shows a search interface with logos for Ontology Engineering Group and BNE. The search bar contains 'Cervantes Saavedra, Miguel de'. Below the search bar, the results are displayed under the heading 'Cervantes Saavedra, Miguel de'. The results include:

- Título: [Visita a las galerías y Nómadas]
- Título: Don Quijote de la Mancha
- Título: El celoso extremo
- Título: Novela
- Título: Leben und Taten des s... Quixote de la Mancha
- Título: Las mejores páginas de estudios y comentarios seguidas de un vocabulario
- Título: La española inglesa
- Título: La cueva de Salamanca

Below the titles is a portrait of Miguel de Cervantes. The page also includes a section titled 'Atributos' with a list of name variants.



Atributos

- tiene nombre de persona:
Cervantes Saavedra, Miguel de
- tiene fechas de persona:
1547-1616
- tiene otra variante de nombre:
Cervantes
- tiene otra variante de nombre:
Cervantes di Saavedra, Michele
- tiene otra variante de nombre:
Cervantes, Miguel de
- tiene otra variante de nombre:
Сервантес Сааведра, Мигель де
- tiene otra variante de nombre:
Θερβάντες, Μιχάλης
- tiene otra variante de nombre:

- How many times you write an URI when developing an application?
 - e.g. var workURI = "http://ifla.ns..../C1001"
- For issuing queries to open SPARQL endpoints opaque URIs are painful.
- **isbd:hasStatementOfResponsibilityRelatingToEdition**
Or
isbd:P1010?

LABELING

Labeling: Initial guidelines for the Multilingual Web

- How to represent labels?

rdfs:label, SKOS, SKOS-XL, Lemon?

- How to structure the content?

Guidelines for the Multilingual Web	Example
Use the singular form for nouns	
Use space as word delimiters, or follow your own language conventions	skos:prefLabel "financial asset"@en
Use upper or lower case according to your own language conventions	rdfs:label "autor"@es rdfs:label "Miguel de Cervantes"@es
Add as many labels as needed for classes and properties	ex:FAOlabel1 rdf:type skosxl:Label; skosxl:literalForm "Food and Agriculture Organization"@en. ex:FAOlabel2 rdf:type skosxl:Label; skosxl:literalForm "FAO"@en

Example: ISBD Cartographic

Preferred Label:	cartographic
Language:	English
URI:	http://iflastandards.info/ns/isbd/terms/contentqualification/type/T1001 (RDF)
Top Concept?:	✓
Status:	Published

Properties

scope note	Uključuje geografske karte, atlase, globuse, reljefne modele i tako dalje.	Serbian	New-Proposed
scope note	Includes maps, atlases, globes, relief models, etc.	English	Published
scope note	Include le carte geografiche, gli atlanti, i globi, i modelli in rilievo, etc.	Italian	Published
scope note	包括地图、地图册、球仪、地形模型等。	Chinese	New-Proposed
scope note	Incluye mapas, atlas, globos, modelos en relieve, etc.	Spanish	Published
scope note	Включва карти, атласи, глобуси, модели на релеф и др.	Bulgarian	New-Proposed
scope note	Ietver kartes, atlantus, globusus, reljefa modejus, u.c.	Latvian	New-Proposed
scope note	Включает в себя карты, атласы, глобусы, модели рельефа и т.п.	Russian	New-Proposed

alternative label	cartográfico	Spanish	New-Proposed
alternative label	cartográfica	Spanish	New-Proposed
alternative label	kartografski	Croatian	New-Proposed
alternative label	kartografska	Croatian	New-Proposed

```
isbd:T1001 lemon:isReferenceOf [ lemon:isSenseOf :cartographic] .  
:cartographic lemon:LexicalEntry ;  
    lemon:form [ lemon:writtenRep "cartográfico"@es ;  
                isocat:grammaticalGender isocat:mASCULINE ] ;  
    lemon:form [ lemon:writtenRep "cartográfica"@es ;  
                isocat:grammaticalGender isocat:fEMININE ] .  
  
:isocat:grammaticalGender rdfs:subPropertyOf lemon:property .
```

Some questions arise:

- What pattern should IFLA use to name/manage these new URIs (:cartographic) ?
- Are we making it more difficult for data consumers (more HTTP requests, etc.)?

THANK YOU!!

Slides are available at **slideshare.net/DanielVilaSuero**

email: **dvila@fi.upm.es**

Twitter: **@dvilasuero**