

DBpedia linkage analysis leveraging on entity semantics

Bc. David Fuchs

Content

DBpedia linkage analysis

- data availability
- results

Functional Requirements for Bibliographic Records (FRBR) on Wikidata

- mapping to Wikidata entities and properties
- consistency within Wikidata

DBpedia links to Wikidata

- annotating DBpedia with FRBR information
- analysis of collected information

DBpedia linkage analysis

- source data
 - links to datasets
 - number of links
- <https://wiki.dbpedia.org/services-resources/interlinking>
- datasets
 - 181 in total
 - 82 provide access to data
- linking triples – subject and object from different datasets

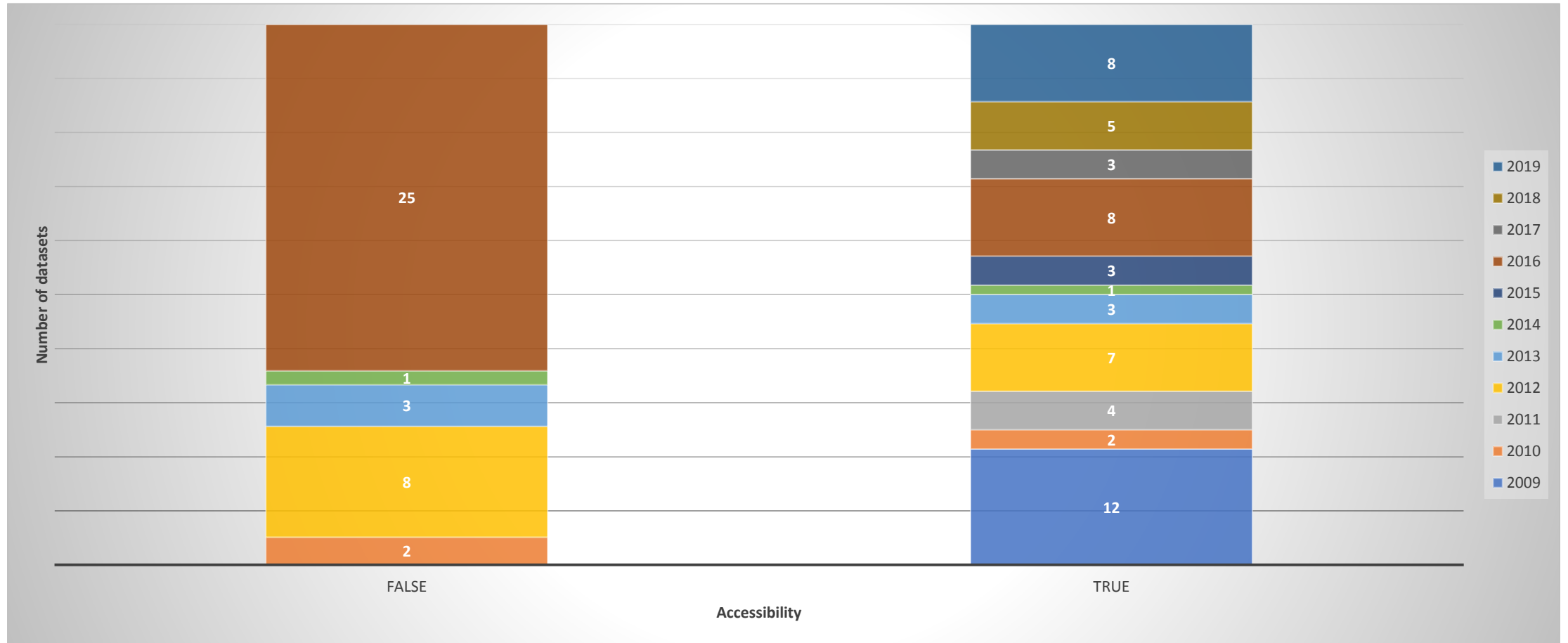
Data Set	Number of Links**
ACM (RKBExplorer)	5
AEMET metereological dataset	82
AGROVOC	993
Airports	9,761
Alpine Ski Racers of Austria	921
Amsterdam Museum	43
BBC Music	23,000
BBC Programmes	23,237
BBC Wildlife Finder	415
BFS LD	261
BibBase	53
Bible Ontology	371
Brazilian Politicians	1,500

DBpedia linkage analysis

- accessibility of datasets
- method used for accessing the dataset
- date of dataset inspection
- alternative access method
- date of last modification of dataset

Data Set	Accessible	Data source type	Date	Alternative access	Last modified
ACM (RKBExplorer)	true	custom web API, SPARQL, dump	15.07.2019		31.12.2009
AEMET metereological dataset	true	SPARQL	15.07.2019		19.09.2011
AGROVOC	false		15.07.2019		
Airports	false		15.07.2019	https://archive.org/details/kasabi	20.07.2012
Alpine Ski Racers of Austria	true	SPARQL, dump	15.07.2019		25.11.2013
Amsterdam Museum	false	SPARQL	15.07.2019		
BBC Music	false		15.07.2019	https://archive.org/details/kasabi	20.07.2012
BBC Programmes	false		15.07.2019	https://archive.org/details/kasabi	20.07.2012
BBC Wildlife Finder	false		15.07.2019	https://archive.org/details/kasabi	20.07.2012
BFS LD	false	SPARQL, dump	15.07.2019		
BibBase	false	SPARQL	15.07.2019		
Bible Ontology	false		15.07.2019		
Brazilian Politicians	false		25.11.2019		

DBpedia linkage analysis



DBpedia linkage analysis

- check existence of triples linking the DBpedia dataset:

ask

where{

 ?subject ?predicate ?object.

 FILTER (strstarts(str(?object),"http://dbpedia.org/") &&
 !(strstarts(str(?subject),"http://dbpedia.org/"))).

}

DBpedia linkage analysis

- DBtune
 - 3153 unique linked entities on DBpedia
 - 32 linked twice (from different entities)
 - dbtune:bohmann_boris
 - dbtune:boris_bohmann
- EARTH

Subject	skos:closeMatch	skos:mappingRelation	skos:semanticRelation	skos:exactMatch
earth:68690	dbr:Angular_resolution	dbr:Angular_resolution	dbr:Angular_resolution	
earth:68690	dbr:Spectral_resolution	dbr:Spectral_resolution	dbr:Spectral_resolution	
earth:24160	dbrc:Data_deficient_species	dbrc:Data_deficient_species	dbrc:Data_deficient_species	dbrc:Data_deficient_species

FRBR on Wikidata

- Where to look for inconsistencies?
 - DBpedia links to Wikidata (through annotating DBpedia resources)
 - DBpedia annotations (different annotators)
 - Wikidata
 - combination of properties
 - combination of a class and property

FRBR on Wikidata

- music
 - composition (Q207628) ~ *Work* ⊆ *Expression*
 - release (Q2031291) ~ *Manifestation*
- books
 - written work (Q47461344) ~ *Work*
 - edition (Q3331189) ~ *Expression* ⊆ *Manifestation*
 - exemplar items (Q53731850) ~ *Item*
 - manuscripts
(Q87167, Q213924, Q1440453, Q834459, Q2217259, Q274076, Q1754581, Q690851, Q284465) ~ *Item*
 - periodicals and collections of works ~ *Work*



FRBR on Wikidata

Type	Title	ID	Data type
Edition	SUDOC editions	P1025	External identifier
Edition	legal citation of this text	P1031	String
Work	Dewey Decimal Classification	P1036	String
Edition	EUL editions	P1084	External identifier
Work	LibraryThing work ID	P1085	External identifier
Edition	illustrator	P110	Item

- properties for types of entities
- descriptions
- examples

FRBR on Wikidata

- inconsistencies in the usage of properties:

```
PREFIX : <https://www.wikidata.org/wiki/Property:>
```

```
ask{
```

```
  ?entity ?work_property [].
```

```
  ?entity ?edition_property [].
```

```
  values ?work_property {
```

```
    # Properties associated with FRBR Work level.
```

```
  }.
```

```
  values ?edition_property {
```

```
    # Properties associated with FRBR Expression  
    # or Manifestation levels.
```

```
  }.
```

```
}
```

- inconsistencies in the combination of type and property:

```
PREFIX : <https://www.wikidata.org/wiki/Property:>
```

```
PREFIX wd: <http://www.wikidata.org/entity/>
```

```
ask{
```

```
  # Class for FRBR Work (Wikidata Written Work)
```

```
  ?entity :P31 wd:Q47461344.
```

```
  ?entity ?edition_property [].
```

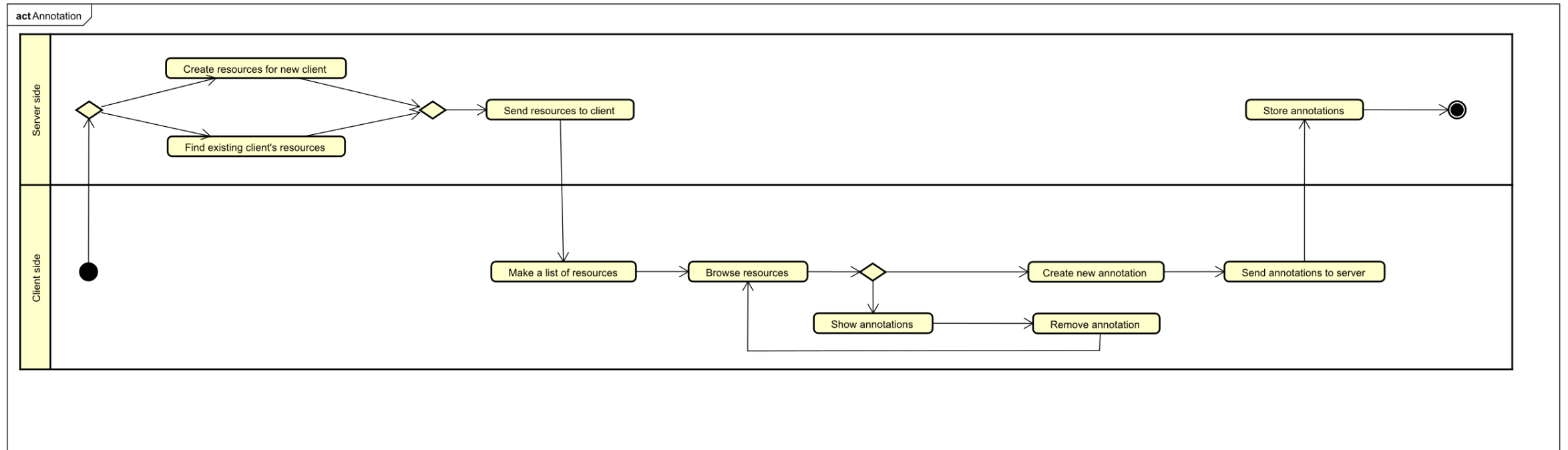
```
  values ?edition_property {
```

```
    # Properties associated with FRBR Expression  
    # or Manifestation levels.
```

```
  }.
```

```
}
```

DBpedia links to Wikidata



DBpedia links to Wikidata

```
PREFIX wdt:
<http://www.wikidata.org/prop/direct/>
PREFIX wd: <http://www.wikidata.org/entity/>
construct {
  ?wikidata_resource wdt:P31 ?wikidata_class.}
WHERE {
  ?wikidata_resource wdt:P31 ?wikidata_class.
  VALUES ?wikidata_class {
    # All Wikidata classes that correspond to
    # FRBR categories
  }.
}
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
construct{
  ?dbpedia_resource owl:sameAs ?wikidata_resource.
}
where{
  ?dbpedia_resource owl:sameAs ?wikidata_resource.
  FILTER (strstarts(str(?dbpedia_resource),"http://dbpedia.org/")).
  FILTER (strstarts(str(?wikidata_resource),"http://www.wikidata.org/")).
}
```

DBpedia links to Wikidata

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX wd: <http://www.wikidata.org/entity/>
construct { ?dbpedia_resource ?dbpedia_property ?dbpedia_value. }
WHERE {
  ?dbpedia_resource owl:sameAs ?wikidata_resource.
  ?wikidata_resource wdt:P31 ?wikidata_class.
  VALUES ?wikidata_class { # Wikidata corresponding to FRBR categories
  }.
  SERVICE <http://dbpedia.org/sparql> { ?dbpedia_resource ?dbpedia_property ?dbpedia_value. }
  FILTER (!isBlank(?dbpedia_resource) && !isBlank(?dbpedia_value)).
  FILTER (strstarts(str(?dbpedia_resource),"http://dbpedia.org/")).
  FILTER (strstarts(str(?wikidata_resource),"http://www.wikidata.org/")).
}
```

DBpedia links to Wikidata: future work

- analysis of collected information
 - <https://github.com/Fuchs-David/Annotator>
 - <https://github.com/propi/rdfrules>
 - head of rule should belong to `<http://vocab.org/frbr/core.html#>`
 - <https://github.com/Fuchs-David/RDFExcelIO> as an alternative
 - analysis and/or preprocessing

Summary



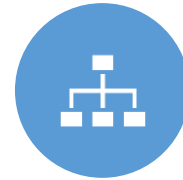
ACCESSIBILITY OF
DATASETS LINKING TO
DBPEDIA



VERIFICATION OF
LINKING TRIPLES
EXISTENCE



ANALYSIS OF DATASETS
WITH VERIFIED LINKING
TRIPLES



INTERNAL
CONSISTENCY OF
WIKIDATA WITHIN
FRBR DOMAIN



ANNOTATION OF
DBPEDIA WITH FRBR
INFORMATION



ACTIONABLE MODEL
BASED ON DBPEDIA
ANNOTATIONS



INTERVIEWS/QUESTIO
NNAIRES ABOUT
DISCOVERED ISSUES

Sources

- Interlinking. *wiki.dbpedia.org* [online]. [Date Accessed: 15-07-2019]. Available at: <https://wiki.dbpedia.org/services-resources/interlinking>
- Wikidata:WikiProject Books. *wikidata.org* [online]. [Date Accessed: 22-11-2019]. Available at: https://www.wikidata.org/wiki/Wikidata:WikiProject_Books
- Wikidata:WikiProject Music. *wikidata.org* [online]. [Date Accessed: 22-11-2019]. Available at: https://www.wikidata.org/wiki/Wikidata:WikiProject_Music