Building Ontologies for Reuse
Lessons Learned from Unit Ontologies

Sirko Schindler$^1$ and Jan Martin Keil$^2$

$^1$DLR Institute of Data Science, Jena, Germany
$^2$Friedrich Schiller University Jena, Germany
Motivation

• Reuse - a core ontology promise
  • Linked Data Cloud

• Different issues hindering reuse of existing ontologies
  • Redundant development

• Units are a common part of many ontologies
  • Multiple ontologies available
  • Reuse has its pitfalls …
Choosing Good IRIs

Identity vs. Equality

Properties
Choosing Good IRIs

Identity vs. Equality

Properties
1a. IRIs should not contain the ontology version.

Dislike: Old versions may disappear from the web, breaking …
   - Resolving of IRIs
   - `owl:imports`

Information: Need references to specific versions? `owl:versionIRI`!
   - Generic IRIs for reuse – always referring to latest version
   - Version-specific ontology-IRIs for provenance, …

http://www.wurvoc.org/vocabularies/om-1.6/Unit_of_measure
http://www.wurvoc.org/vocabularies/om-1.8/Unit_of_measure
1b. IRIs should not be too long.

👎 Cluttering (text-)editors

👎 Consuming valuable resources in restricted environments
  - IoT RDF streams, …

👍 Limit the length of your IRIs

👍 Generic local-names
  - `wd:Q67006310` “2nd International Workshop on Bad Or Good Ontology” (Wikidata.org)
  - `bfo:BFO_0000006` “spatial region” (BFO – OBOFoundry.org)
  - …

1c. IRIs of large resource collections should not contain natural language.

.fixing typos breaks references

Language dependent

Generic local-names and attached labels

- \textit{wd:Q13298} (Wikidata.org)
  - “Graz”@de
  - “Γκρατς”@el
  - “Graz”@en
  - “گراتس”@fa
  - ...

\textit{http://purl.oclc.org/NET/muo/ucum/unit/pressure/pound-per-square-inch}
1d. Prefixes should not refer to multiple namespaces.

- In modularized ontologies

Possible Namespace mix-ups

Be consistent

Use namespace lookups like [https://prefix.cc](https://prefix.cc)

```xml
<!ENTITY comp "http://sweet.jpl.nasa.gov/2.3/reprSciComponent.owl#">
<!ENTITY comp "http://sweet.jpl.nasa.gov/2.3/statePhysical.owl#">
<!ENTITY comp "http://sweet.jpl.nasa.gov/2.3/matrCompound.owl#">
```
1e. Namespaces should not be referred by multiple prefixes.

- In modularized ontologies

👎 Namespace confusion

👍 Be consistent

👍 Check the results in generated ontologies

http://sweet.jpl.nasa.gov/2.3/propOrdinal.owl

<!ENTITY screla "http://sweet.jpl.nasa.gov/2.3/relaSci.owl#">

http://sweet.jpl.nasa.gov/2.3/propEnergyFlux.owl

<!ENTITY screla2 "http://sweet.jpl.nasa.gov/2.3/relaSci.owl#"
1f. Namespaces should not omit the hash.

👎 Fine with XML, but breaks in Turtle or SPARQL
  • Hash # is used for comments
👎 Different prefixes/namespaces in reuse

```sparql
PREFIX screla: <http://sweet.jpl.nasa.gov/2.3/relaSci.owl>
PREFIX units: <http://sweet.jpl.nasa.gov/2.3/reprSciUnits.owl>

SELECT ?quantity
WHERE {
  # get all quantities that use lumen
  ?quantity screla:#hasDefaultUnit units:#lumen .
}
```

http://sweet.jpl.nasa.gov/2.3/propEnergyFlux.owl

<!ENTITY screla "http://sweet.jpl.nasa.gov/2.3/relaSci.owl"
Choosing Good IRIs

Identity vs. Equality

Properties
2a. Do not confuse equivalency and identity.
2b. Be aware of alleged synonyms.

Dislike Possible oversimplifications

Dislike Mathematically equivalent – not identical
  • Conversion factor of one

Inform owl:sameAs means identical
  • Both IRIs exchangeable in all contexts!
  • Same for skos:altLabel etc.

Inform Other levels of “sameness”
  • E.g., skos:broadMatch, skos:narrowMatch, skos:closeMatch, skos:exactMatch, ...

“cubic metre per square metre” vs “metre”
“litre” vs “cubic decimetre”
2c. Know the exceptions.

- Things are actually renamed sometimes.

Two IRIs
- Linked by `owl:sameAs`

One IRI
- Multiple labels, e.g., `skos:altLabel`

---

The gon (or grad, where grad is an alternative name for the gon) is an alternative unit of plane angle to the degree, defined as \((\pi/200)\) rad.

Choosing Good IRIs

Identity vs. Equality

Properties
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions

(a) OBOE

(b) OM

(c) QU
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions

👎 Unclear models

👍 Prevent misinterpretation

👍 Add descriptions
3a. Properties should be modeled resilient against misinterpretation.

- Different models for unit conversions

**Unclear models**

**Prevent misinterpretation**

**Add descriptions**

(a) OBOE

(b) OM
3b. Dependent properties should be encapsulated into distinct resources.

- Possible only once per resource
- Losing the dependency between values

- Add a distinct resource
  - Preserve the dependency
  - Allow for multiple definitions
Summary

Issues encountered while reusing unit ontologies

- IRI design
- Identity vs. Equality
- Property Modeling

Hints to prevent them
Besides tool support …

👍 Reusability testing in ontology development
  • Usability testing in user interface design

👍 Have others use your ontology before publishing.
Image Credits

- Motivation: https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg
- 1a: https://commons.wikimedia.org/wiki/File:Human_evolution_scheme.svg
- 1b: https://pixabay.com/photos/letters-letter-wooden-alphabet-4466230/
- 1c: data @ Wikidata
- 1d: http://www.freestockphotos.biz/stockphoto/8213
- 1e: https://pxhere.com/en/photo/984500
- 1f: https://pixabay.com/illustrations/faulty-system-chair-broken-team-4295549/
- 2a: https://commons.wikimedia.org/wiki/File:Green_pea_pod_8872.jpg
- 2b: https://commons.wikimedia.org/wiki/File:The_McClure_Twins.jpg
- Summary: https://pixabay.com/illustrations/tropical-summer-sunset-beach-1651426/
- Reusability testing: https://knowyourmeme.com/memes/forwards-from-grandma