FAIR principles application to Semantic Web and ontologies

Towards Implementations of Materials and Manufacturing Commons Workshop

María Poveda-Villalón, Ontology Engineering Group
ETSI Informáticos
Universidad Politécnica de Madrid, Spain

mpoveda@fi.upm.es  @MariaPovedaV  5th April 2023  Berlin, Germany
Introduction

Linked Data principles

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs so that they can discover more things.

Adoption:
- EOSC interoperability framework
- Research Data Alliance

https://www.nature.com/articles/sdata201618
There is a clear movement towards expanding the application of the FAIR principles beyond research data [EOSC Interoperability Framework]

Ontologies are often the result of research activities or fundamental components in many research areas

Some initiatives (FAIRsFAIR EU Project recommendations, GO-FAIR implementation network GO-INTER, RDA Vocabulary Services Interest Group, “Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web”...)

How do these works fit with the Ontology Engineering community and the Semantic Web practices?

17 recommendations, related to one or more FAIR principles related to:

- **GUPRIs** (Global Unique Persistent and Resolvable Identifier)
- (minimum) **metadata** including provenance, license, etc.
- **Semantic repositories**
  - API
  - Cross access
  - Secure protocols
- Use standards (languages, vocabularies)
- Mappings (between artefacts, to foundational ontologies)

https://doi.org/10.5281/zenodo.3707985
10 guidelines for publishing FAIR **ontologies** and **vocabularies** related to:
- **Accessible** and **permanent** ontology URIs
- Generation of reusable **documentation** (metadata and human oriented)
- **Publication** of ontologies on the Web (formats, findable)


“Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web”
Publish your vocabulary on the Web at a stable URI with an open license.

Provide human-readable documentation and basic metadata such as creator, publisher, date of creation, last modification, version number.

Provide labels and descriptions, if possible in several languages, to make your vocabulary usable in multiple linguistic scopes.

Make your vocabulary available via its namespace URI, both as a formal file and human-readable documentation, using content negotiation.

Link to other vocabularies by re-using elements rather than re-inventing.

There is dereferenceable human-readable information about the used vocabulary.

The information is available as machine-readable explicit axiomatization of the vocabulary.

The vocabulary is linked to other vocabularies.

Metadata about the vocabulary is available (in a dereferencable and machine-readable form).

The vocabulary is linked to by other vocabularies.


“Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web”

5-star vocabularies
Vatant, Bernard 2012

5-star vocabularies
SWJ 2014
## Towards FAIR Ontologies – To be Findable

<table>
<thead>
<tr>
<th>Keep from SW</th>
<th>Needs</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1: URIs</td>
<td></td>
<td>Persistence</td>
</tr>
<tr>
<td>F2: Metadata included in the ontology</td>
<td>Minimum metadata, technical guidelines</td>
<td>Metadata as a separate object, third-party certifier</td>
</tr>
<tr>
<td>F3: DCAT2</td>
<td>Federation model, SAODs</td>
<td></td>
</tr>
</tbody>
</table>

- **F1:** (meta)data are assigned a globally unique and persistent identifier
- **F2:** data are described with rich metadata (defined by R1 below)
- **F3:** metadata clearly and explicitly include the identifier of the data it describes
- **F4:** (meta)data are registered or indexed in a searchable resource
Towards FAIR Ontologies – To be Accessible

<table>
<thead>
<tr>
<th>Keep from SW</th>
<th>Needs</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>URIs</td>
<td>Persistence</td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td>Minimum metadata, technical guidelines</td>
</tr>
<tr>
<td>F3</td>
<td>Metadata included in the ontology</td>
<td>Metadata as a separate object, third-party certifier</td>
</tr>
<tr>
<td>F4</td>
<td>DCAT2</td>
<td>Federation model, SAODs</td>
</tr>
<tr>
<td>A1, A1.1, A1.2</td>
<td>HTTP and HTTPS</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td>Preservation policies</td>
</tr>
</tbody>
</table>

- A1: (meta)data are retrievable by their identifier using a standardized communications protocol
- A1.1: the protocol is open, free, and universally implementable
- A1.2: the protocol allows for an authentication and authorization procedure, where necessary
- A2: metadata are accessible, even when the data are no longer available
# Towards FAIR Ontologies – To be Interoperable

<table>
<thead>
<tr>
<th></th>
<th>Keep from SW</th>
<th>Needs</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>URIs</td>
<td>Minimum metadata, technical guidelines</td>
<td>Persistence</td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>Metadata included in the ontology</td>
<td>Federation model, SAODs</td>
<td>Metadata as a separate object, third-party certifier</td>
</tr>
<tr>
<td>F4</td>
<td>DCAT2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1, A1.1, A1.2</td>
<td>HTTP and HTTPS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I1</td>
<td>KR languages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I2</td>
<td>Methods to reuse ontologies</td>
<td>Indicators</td>
<td>Not force to reuse FAIR vocabularies</td>
</tr>
<tr>
<td>I3</td>
<td>Mechanisms to reference ontologies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **I1**: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- **I2**: (meta)data use vocabularies that follow FAIR principles
- **I3**: (meta)data include qualified references to other (meta)data
Towards FAIR Ontologies – To be Reusable

<table>
<thead>
<tr>
<th>Keep from SW</th>
<th>Needs</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>URIs</td>
<td>Minimum metadata, technical guidelines</td>
</tr>
<tr>
<td>F2</td>
<td>Metadata included in the ontology</td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>DCAT2</td>
<td>Federation model, SAODs</td>
</tr>
<tr>
<td>F4</td>
<td>HTTP and HTTPS</td>
<td></td>
</tr>
<tr>
<td>A1, A1.1, A1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I1</td>
<td>R1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation</td>
<td></td>
</tr>
<tr>
<td>I2</td>
<td>R1.1: (meta)data are released with a clear and accessible data usage license</td>
<td></td>
</tr>
<tr>
<td>I3</td>
<td>R1.2: (meta)data are associated with detailed provenance</td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td></td>
<td>Best practices for document and communicate ontologies</td>
</tr>
<tr>
<td>R1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R1.2</td>
<td>Link to the license URI or RDF description of it</td>
<td>PROV-O</td>
</tr>
<tr>
<td>R1.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Best practices for document and communicate ontologies
## Towards FAIR Ontologies

- **Keep from SW**
  - F1: URIs
  - F2:
  - F3: Metadata included in the ontology
  - F4: DCAT2
  - A1, A1.1, A1.2: HTTP and HTTPS

- **Needs**
  - Minimum metadata, technical guidelines
  - Metadata as a separate object, third-party certifier
  - Federation model, SAODs

- **Discussion**
  - Persistence
  - Preservation
  - Good news here!
  - Indicators
  - Not force to reuse FAIR vocabularies

- **Indicators**
  - Best practices for document and communicate ontologies

- **Link to the license URI or RDF description of it**
- **PROV-O**

- **Community standards**
Classical/Core Workflow for building ontologies

Legend
- Actors
- Activity
- Artefacts

Ontology requirements specification → Ontology implementation → Ontology publication → Ontology maintenance

FAIR assessment usually happens here

http://lot.linkeddata.es/
First steps to support the validation

- Validation service inspired by OOPS! (OntOlogy Pitfall Scanner)
- Designed to guide users
  - Tests have an explanation
  - Tests indicate potential errors
- Practical
  - Based on years of ontology engineering practices at UPM
- Aligned to FAIR

https://w3id.org/foops/

Slide taken from “FOOPS! An Ontology Pitfall Scanner for the FAIR principles. Dbpedia day” by Daniel Garijo
### FOOPS!: Getting the full report

<table>
<thead>
<tr>
<th>Ontology metadata summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAIRness coverage by category</td>
</tr>
<tr>
<td>FAIRness overall score.</td>
</tr>
<tr>
<td>Note: this may be a quality indicator, but there is no defined threshold for FAIRness.</td>
</tr>
</tbody>
</table>

#### FAIR Category

<table>
<thead>
<tr>
<th>Check</th>
<th>Check coverage</th>
<th>Check description</th>
<th>Check explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Findable</td>
<td>Accessible</td>
<td>Interoperable</td>
<td>RDF1: RDF Availability</td>
</tr>
<tr>
<td>I1: (meta)data use a formal, accessible, shared, and broadly usable language for knowledge representation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC1: Vocabulary reuse (metadata)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC2: Vocabulary reuse</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Slide taken from “FOOPS! An Ontology Pitfall Scanner for the FAIR principles. Dbpedia day” by Daniel Garijo
Classical/Core Workflow for building ontologies

Legend
- Actors
- Activity
- Artefacts

output
- activity flow
- (input) Artefact reference

Thinking in advance

http://lot.linkeddata.es/
Generate reusable documentation
Human-readable documentation, including good ontology visualization

https://chowlk.linkeddata.es/

Adapted from Paola Espinoza Arias
Ontology Implementation

Generate reusable documentation

Ontology metadata to describe ontologies

Adapted from Paola Espinoza Arias
Ontology Implementation

Generate reusable documentation

Ontology metadata to describe terms

<table>
<thead>
<tr>
<th>Property name</th>
<th>Annotation Property</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>rdfs:label</td>
<td>Readability</td>
</tr>
<tr>
<td>Definition</td>
<td>rdfs:comment</td>
<td>Understanding</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Property name</th>
<th>Annotation Property</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>vanн:example</td>
<td>Understanding</td>
</tr>
<tr>
<td>Status</td>
<td>sw:term_status</td>
<td>Understanding</td>
</tr>
<tr>
<td>Rationale</td>
<td>vaem:rationale</td>
<td>Understanding</td>
</tr>
<tr>
<td>Source</td>
<td>dterms:source</td>
<td>Provenance</td>
</tr>
</tbody>
</table>

Adapted from Paola Espinoza Arias
Design accessible ontology URIs:

1. Ontology name and prefix
2. Hash or slash
3. Meaningful or opaque
4. Ontology versioning
5. Permanent URIs

Example

1. name: SAREF extension for Smart Cities, prefix: saref4city
2. https://w3id.org/def/saref4city#
3. https://w3id.org/def/saref4city#AdministrativeArea
4. <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">1.0.0</owl:versionInfo>
5. https://w3id.org

Adapted from Paola Espinoza Arias
Ontology implementation

Reuse (FAIR) ontologies
Maximize interoperability
Ontology publication

Generate reusable documentation

**Human-readable** documentation, including good ontology visualization, examples of use, queries, etc.
Publish the ontology on the Web

Provide several interoperable **formats**

URI

- Own URI
- purl, w3id, etc.
- Content negotiation

https://www.w3.org/TR/swbp-vocab-pub/
Publish the ontology on the Web

Make the ontology findable

Ontology publication

Ontology

Publication

- Ont. Devel.
- Ont. Devel. Experts
- Ont. Devel.

Ontology requirements specification

Ontology implementation

Propose release candidate

Documentation

Online publication

Ontology maintenance

ORSD

Ontology

Online ontology

HTML documentation

Online access (content negotiation)

Issue tracker

ontobee

BioPortal

FAIRsharing

Industry Portal

VCINITY

prefix.cc

https://prefix.cc/
Conclusions

- **Metadata** is one of requirements to produce FAIR ontologies.
  - But **not** the **only** one

- Adopt **existing practices** and **technologies**

- Think about FAIR principles at all stages of the development
  - But note that FAIR **doesn’t** look at the resource **quality**!
Thank you very much for your attention!

Contact: mpoveda@fi.upm.es

Questions?