| | |
|---|---|
| Project Title | Fostering FAIR Data Practices in Europe |
| Project Acronym | FAIRsFAIR |
| Grant Agreement No | 831558 |
| Instrument | H2020-INFRAEOSC-2018-4 |
| Topic | INFRAEOSC-05-2018-2019 Support to the EOSC Governance |
| Start Date of Project | 1st March 2019 |
| Duration of Project | 36 months |
| Project Website | [www.fairsfair.eu](http://www.fairsfair.eu) |

# D2.8 FAIR Semantics Recommendations Third Iteration

| | |
|---|---|
| Work Package | WP2 |
| Lead Author (Org) | Yann Le Franc (eSDF) |
| Contributing Author(s) (Org) | Luiz Bonino (LUMC - DTL), Hanna Koivula (CSC), Jessica Parland-von Essen (CSC), Robert Pergl (eSDF) |
| Due Date | 28.02.2022 |
| Date | 25.02.2022 |
| Version | 1.0 Draft not yet approved by the European Commission |
| DOI | 10.5281/zenodo.6276576 |

Dissemination Level

- [ ] PU: Public
- [ ] PP: Restricted to other programme participants (including the Commission)
- [ ] RE: Restricted to a group specified by the consortium (including the Commission)
- [ ] CO: Confidential, only for members of the consortium (including the Commission)

## Abstract

This document is the third and final iteration of recommendations for making semantic artefacts FAIR. These recommendations result from continuous discussions with semantic experts from multiple communities. Our previous work included 17 preliminary recommendations related to one or more of the FAIR principles, and 10 best practice recommendations on semantic artefacts. These recommendations were last published as Deliverable 2.5 and have now gone through minor revisions. The work has been published on GitHub and we used GitHub's issue tracking feature to allow the community to comment on the recommendations and best practices. The work presented in this version relates to the Best practices, the proposition for an initial service architecture to support FAIR Semantics, a first version of a community-driven minimum metadata schema for describing the Semantic Artefacts and discussing the future work around the recommendation and FAIR semantics.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRSFAIR
Fostering Fair Data Practices in Europe

## Versioning and Contribution History

| Version | Date | Authors | Notes |
|---|---|---|---|
| 0.5 | | Yann Le Franc | Revision and simple improvements to 2nd iteration recommendations and best practices |
| 0.6 | | | Revision Executive Summary, Introduction |
| 0.7 | | Hanna Koivula | Addition of Best Practices analysis |
| 0.8 | | Yann Le Franc, Robert Pergl, Luiz Bonino | Addition of section 4 |
| 0.9 | 21/02/2022 | Robert Huber and Eva Mendez | Internal review |
| 1.0 | 24/02/2022 | Yann Le Franc | Revision and response to internal review |

## Disclaimer

FAIRsFAIR has received funding from the European Commission's Horizon 2020 research and innovation programme under the Grant Agreement no. 831558 The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## Abbreviations and Acronyms

| | |
|---|---|
| API | Application Programming Interface |
| BFO | Basic Formal Ontology |
| BP | Best Practice |
| BP-Rec | Best Practice Recommendation |
| CODATA | Committee on Data of the International Science Council |
| DCMI | Dublin Core Metadata Initiative |
| DCAT | Data Catalogue Vocabulary |
| DOI | Digital Object Identifier |
| DOLCE | Descriptive Ontology for Linguistic and Cognitive Engineering |
| DOOR | Descriptive Ontology of Ontology Relations |
| ELIXIR | ELIXIR the European life-sciences Infrastructure for biological Information |
| EMMO | European Materials Modelling Ontology |
| EOSC | European Open Science Cloud |
| ESFRI | European Strategy Forum on Research Infrastructures |
| FAIR | Findable, Interoperable, Accessible and Reusable |
| FDP | FAIR Data Point |
| FDMM | FAIR Data Maturity Model |
| FOAF | Friend Of A Friend |
| GUPRI | Globally Unique, Persistent and Resolvable Identifier |
| HTTP | Hypertext Transfer Protocol |
| IOF | Industrial Ontology Foundry |
| IRI | Internationalised Resource Identifier |
| JSON-LD | JavaScript Object Notation for Linked Data |
| KOS | Knowledge Organisation System |
| LD | Linked Data |
| LOV | Linked Open Vocabularies |
| MIREOT | Minimum Information to Reference an External OnTology |
| MIRO | Minimum Information for the Reporting of an Ontology |
| MOD | Metadata for Ontology Description and publication |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| | |
|---|---|
| NERC | Natural Environment Research Council (of UK) |
| NKOS | Networked Knowledge Organisation Systems |
| OBO | Open Biological and Biomedical Ontology |
| ODRL | Open Digital Rights Language |
| OP | Object Property |
| OMV | Ontology Metadata Vocabulary |
| OWL | Web Ontology Language |
| P-Rec | Preliminary Recommendations |
| PID | Persistent Identifier |
| PURL | Persistent Uniform Resource Locator |
| RDA | Research Data Alliance |
| RDA VSIG | Research Data Alliance Vocabulary Services Interest Group |
| RDF | Resource Description Framework |
| RDFS | Resource Description Framework Schema |
| Rec | Recommendation (See. P-Rec, BP-Rec) |
| RIF-CS | Registry Interchange Format - Collections and Services |
| SHACL | Shapes Constraint Language |
| SHARC | SHARing Rewards and Credit |
| SKOS | Simple Knowledge Organisation System |
| SSSOM | Simple Standard for Sharing Ontology Mappings |
| SPARQL | SPARQL Protocol and RDF Query Language |
| TFiR | Turning FAIR into Reality |
| UFO | Unified Foundational Ontology |
| URI | Uniform Resource Identifier |
| URL | Uniform Resource Locator |
| W3C | World Wide Web Consortium |
| XML | Extensible Markup Language |

## Definition of Important Terms



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| | |
|---|---|
| Controlled vocabulary | A controlled vocabulary is a normalised, restricted list of terms for a specific use or context. Thesauri and taxonomies are types of controlled vocabularies, but not all controlled vocabularies are thesauri or taxonomies. |
| Glossary | A glossary is an alphabetical list of terms in a particular domain of knowledge with the definitions for those terms. |
| Ontology | An ontology is a formal version of a thesaurus where relations are described using a formal system such as Description Logic (DL) to mathematically classify individuals of classes and properties |
| Semantic artefact | A semantic artefact is defined in this work as a machine-actionable and -readable formalisation of a conceptualisation, enabling sharing and reuse by humans and machines. These artefacts may have a broad range of formalisation, from loose sets of terms, taxonomies, thesauri to higher-order logics. Moreover, semantic artefacts are serialised using a variety of digital representation formats, e.g., RDF Turtle, and OWL, using XML (RDF) and JSON-LD. |
| Semantic Registry | A semantic registry is a catalogue that contains metadata about semantic artefacts. |
| Semantic Repository | A semantic repository is defined in this recommendation as a service that stores and offers access to both the metadata of semantic artefacts and their content, i.e. offers search and access to get individual terms (including their metadata) both for humans and for machines. |
| Taxonomy | A taxonomy is a controlled vocabulary with a hierarchical structure used to classify things or concepts. Terms within a taxonomy have relations to other terms (parent/broader term, child/narrower term). |
| Term/class/concept | A term/class/concept is an individual element with a unique semantic interpretation, represented with a unique identifier. |
| Thesaurus | A thesaurus is essentially a controlled vocabulary following a standard structure, where all terms have relationships of three kinds to each other: hierarchical (broader term/narrower term), associative (related term), and equivalent (use/used for or see/ seen from). Some terms in thesauri might have additional explanatory notes, such as scope notes (brief explanations about the coverage of the term or of how it should be used in indexing) or history notes. Thesauri are defined in the ISO 25964[1] |

[1] http://www.niso.org/schemas/iso25964



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## Executive Summary

Semantic artefacts (i.e. controlled vocabularies, ontologies, thesauri, and other knowledge organisation systems) are key building blocks for the implementation of the FAIR principles, specifically as emphasised in the Interoperability principle I2 "*(Meta)data use vocabularies that follow FAIR principles*". However, most of these artefacts are actually not FAIR themselves.

The main objective of our work within the Task "FAIR Semantics" of the FAIRsFAIR project is to support the creation of a federated semantic space by harmonising practices in the development and usage of semantics in representing and describing information and knowledge. For this purpose, we are working to establish guidelines for practitioners, repositories, the community, and any related stakeholders. To ground these recommendations in reality, we are collecting recommendations and practical information from practitioners through an open consultation and dedicated workshops, and we are reusing/ referring to existing recommendations built by different communities of practice.

This document summarises a third iteration of our work focused on the practical aspects of the community-driven recommendations for making semantic artefacts FAIR. These recommendations and Best practices have been discussed extensively with various communities and reached a certain level of consensus on the second iteration of the recommendations, published in D2.5 in the beginning of 2021. Within this document, we are presenting our initial steps toward the practical aspects of the implementation of the recommendations and in particular the creation of a minimal metadata schema for describing Semantic artefacts and a proposal of service architecture to ease the work with Semantic artefacts. We are also discussing the necessary alignment of existing Best practices to reduce the confusion created by the existence of various best practices leading to incompatibility and interoperability issues.

The FAIR Semantics team has also been engaged in active participation in a number of international and pluri-disciplinary initiatives such as RDA, CODATA, and GO FAIR in order to foster grassroots engagement in the recommendation development and ensure that the output delivered is aligned to the needs of its stakeholders. In the previous version of the recommendations, feedback and inputs have been collected from an expert group on 'FAIR Semantic Repositories' leading to updates of the recommendations. In this version, we are leveraging the discussion which occurred within the Minimum Metadata Task Group led by Clement Jonquet.

In the final section of this document we are presenting the status of the recommendation uptake and we discuss the future of the recommendations and in particular its strong ties with the RDA Vocabulary and Semantic Service Interest Group (RDA VSSIG).



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

# Table of Contents





FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

# 1. Introduction and Scope

The FAIR principles are a set of technology and domain-agnostic guidelines to make digital assets Findable, Accessible, Interoperable, and Reusable, defined originally by Wilkinson et al. (2016) in the context of research data. Semantic artefacts (i.e. controlled vocabularies, thesauri, ontologies, etc.) are machine readable models of knowledge. They are used to facilitate the extraction and representation of knowledge within datasets using annotations or assertions. These annotations and assertions enable discovery, interoperability, integration and data retrieval. Both, artefacts and services that support and offer them, play a major role in the implementation of the FAIR principles (particularly principle 2 of Interoperability (I2): (Meta)data - use vocabularies that follow the FAIR principles) and in building FAIR Scientific Knowledge Graphs, in order to express and link scientific contributions and related artefact in a semantically rich FAIR graphical model. This role has been acknowledged by the European Commission Expert Group on FAIR Data in Recommendation 7 within their final report and action plan "Turning FAIR into reality" (European Commission Expert Group on FAIR Data, 2018, p.42):

> *"Support semantic technologies - Semantic technologies are essential for the interoperability and need to be developed, expanded and applied both within and across disciplines"*

According to the expert group on FAIR Data, semantic artefacts and registries have been developed within almost all scientific disciplines. However, semantic artefacts have often been built using different formats (SKOS, XML, RDF, OWL), different levels of complexity/ expressiveness (codelists, reference data, controlled vocabularies, taxonomies, thesauri, ontologies, formal ontologies, etc.) and are scattered on the web. Indeed, in many cases, semantic artefacts are not interoperable and not easy to find and are therefore accessible only to the community of practice within which they were developed, which clearly hampers their reuse (Goldfarb and Le Franc, 2017). The emergence of semantic registries such as BARTOC[2], FAIRsharing[3] and repositories and semantic repositories, such as Bioportal[4] (Whetzel et al., 2011), EBI-OLS[5] (Jupp and al., 2015), Ontobee[6] (Ong et al., 2017), Research Vocabularies Australia[7], the NERC Vocabulary Service[8], Linked Open Vocabulary[9] and others provide means to improve discoverability and enable reusability. The importance of such semantic registries/repositories and the issue of the findability of semantic artefacts is already being

---

[2] BARTOC https://bartoc.org/
[3] FAIRsharing https://fairsharing.org/
[4] Bioportal https://bioportal.bioontology.org/
[5] Ontology Lookup Service https://www.ebi.ac.uk/ols/index
[6] Ontobee: http://www.ontobee.org/
[7] Research Vocabularies Australia https://ardc.edu.au/services/research-vocabularies-australia/
[8] NERC Vocabulary Server https://www.bodc.ac.uk/resources/products/web_services/vocab/
[9] LOV https://lov.linkeddata.es/dataset/lov/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

worked on by various groups such as the Networked Knowledge Organisation System (NKOS)[10], the DCMI/NKOS Interest Group[11] and discussed by d'Aquin & Noy (d'Aquin and Noy, 2012). These services have been identified as key elements for the EOSC Interoperability Framework (Corcho et al., 2021). Despite these existing changes, a large number of semantic resources (i.e. artefacts and repositories) do not comply with most of the FAIR principles.

Semantic Web technologies and standards were built to connect and add meaning to data silos in order to create a web of data next to a web of documents as the current World Wide Web. Unfortunately, in the past decades, the isolated development of semantic artefacts and the lack of common practices to foster interoperability and reusability of semantic artefacts lead to the creation of semantic silos. There is therefore a clear need for a harmonised framework to build, share, publish and reuse semantic artefacts which will provide a harmonised semantic landscape easing reuse and integration for practitioners.

The main goal of **task 2.2 "FAIR Semantics''** is to build such harmonised framework by proposing a set of recommendations and good practices that enable domain specific specialist and data professionals to design FAIR semantic artefacts from the start and therefore de facto supporting the usage of semantics in the FAIRification of data, cross-disciplinary semantic interoperability and the creation of FAIR Scientific Knowledge Graphs. The current situation is characterised by a lack of communication and cross fertilisation between the semantic web and knowledge (ontology) engineering practitioners across various domains of application. To overcome this situation, we are proposing general recommendations that could be applied by all domains of knowledge which are aligned with the individual FAIR principles. To create such generic recommendations, we collected inputs at the grassroots level with the support of as many experts as possible. This will foster the integration of the various existing approaches and support adoption of the recommendations.

Our approach relies on establishing a platform for discussion and collaboration between all stakeholders, to propose a common approach to define recommendations for FAIR Semantics and to promote existing domain-specific efforts, such as OBO foundry[12] for the biomedical domain or the Industry Ontology Foundry[13]. For this purpose, we organised dedicated workshops to gather a large and diverse audience. The initial version of these recommendations were proposed by experts during our first brainstorming session organised as a workshop co-located with RDA Plenary 14 (2019) in Helsinki, and then over the course of 2020 they have been refined and adapted based on stakeholder feedback which culminated in a half-day evaluation workshop in October 2020. Following the publication of this report the new release of FAIR Semantics recommendations will once again be disseminated to the communities to gather feedback.

[10] NKOS: https://nkos.slis.kent.edu/
[11] DCMI/NKOS Interest Group https://dublincore.org/groups/nkos/
[12] OBO Foundry http://www.obofoundry.org/
[13] Industrial Ontology Foundry https://www.industrialontologies.org



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## 1.1. Defining Semantic Artefacts

Initially, we were considering using the common term "ontology" to encompass the different types of semantic models. However, during our discussions both within the project and with our colleagues, we realised that the term "ontology" had different meanings for different communities of practice. This ambiguity of the concept "ontology" has been discussed largely in the scientific literature, for example by Guarino et al. (Guarino et al., 2009) and is still debated (see Neuhaus, 2017).

The original definition has been given by Gruber in 1993: "An ontology is an explicit specification of a conceptualization" (Gruber, 1993). In this context, *"A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose."* Based on these two key definitions, we can consider ontologies as semantic models of a part of the real world.

Due to the problem of ambiguities with the use of the term "ontology", we decide to distance ourselves from this debate by proposing and using a more generic umbrella term: **Semantic Artefact**. Semantic Artefact is defined here as a machine-actionable and -readable formalisation of a conceptualisation enabling sharing and reuse by humans and machines. These artefacts may have a broad range of formalisation, from loose set of terms, taxonomies, thesauri to higher-order logics, and include the concepts/terms/classes constituting these. Moreover, semantic artefacts are serialised using a variety of digital representation formats, e.g., RDF Turtle, OWL-RDF, XML, JSON-LD. In current practices, these artefacts share a common structure encapsulating its metadata, the data i.e. the semantic artefact content comprising of concepts/terms/classes and relations among them, and their (artefact's content) associated metadata (see fig. 1).



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

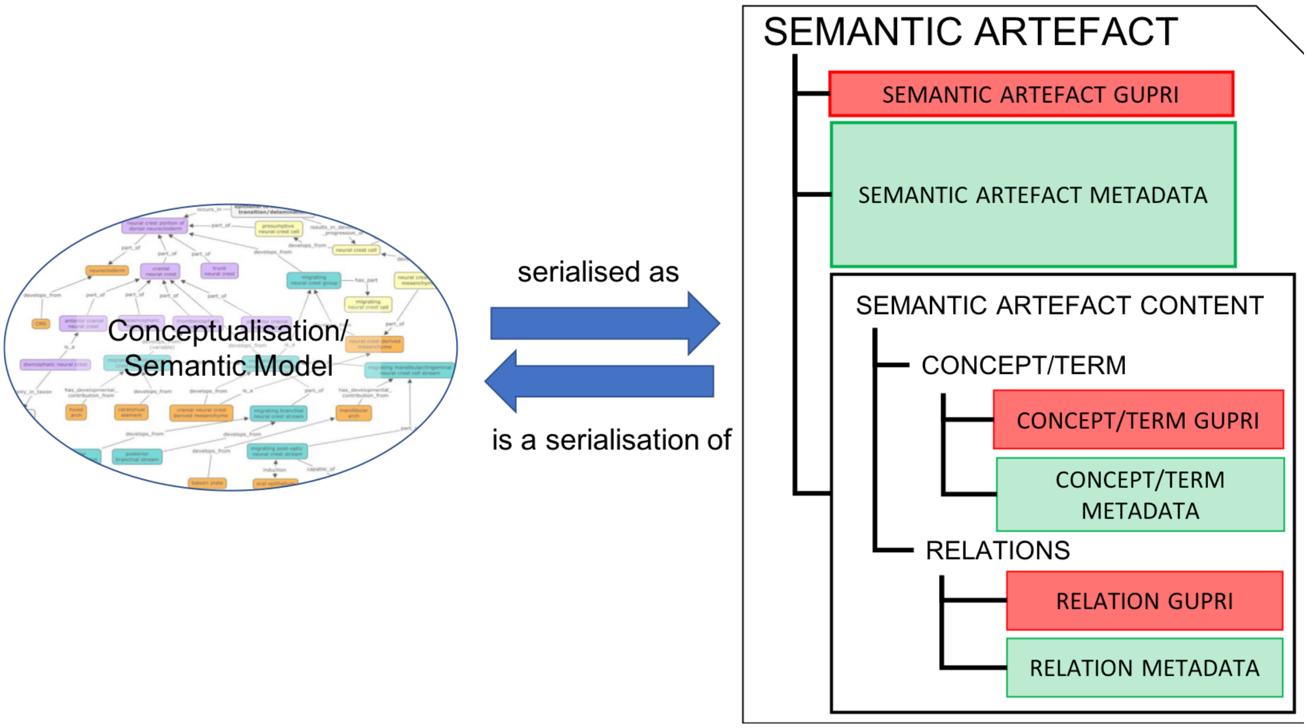


*Figure 1: Common structure of semantic artefacts*

*Semantic artefacts are often structured text files. They have a common structure encapsulating a GUPRI (Globally Unique, Persistent and Resolvable Identifier) for the semantic artefact, the metadata describing the semantic artefact and the semantic artefact content i.e. concept/term and relations. Both are also encapsulating a GUPRI and associated metadata.*

As we mentioned previously, Semantic Artefacts are created in different formats and at different levels of complexity/expressiveness. To classify Semantic Artefacts based on their complexity, an "ontology spectrum" has been proposed and can be used to identify the different types of semantic artefacts and associate them with common formats used to serialise these models (Obrst, 2003, 2010). In fig. 2, we are presenting a simplified version of the spectrum which represents the different types of Semantic Artefacts along a semantic strength axis (ranging from weak semantics to strong semantics). The classification starts with simple lists of terms/concepts (code list, glossary, catalogue ID, controlled vocabulary). These lists of terms/concepts are the simplest building block of semantics and provide a minimal set of information for each item such as a definition, context information and provenance information without relations of any kind.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

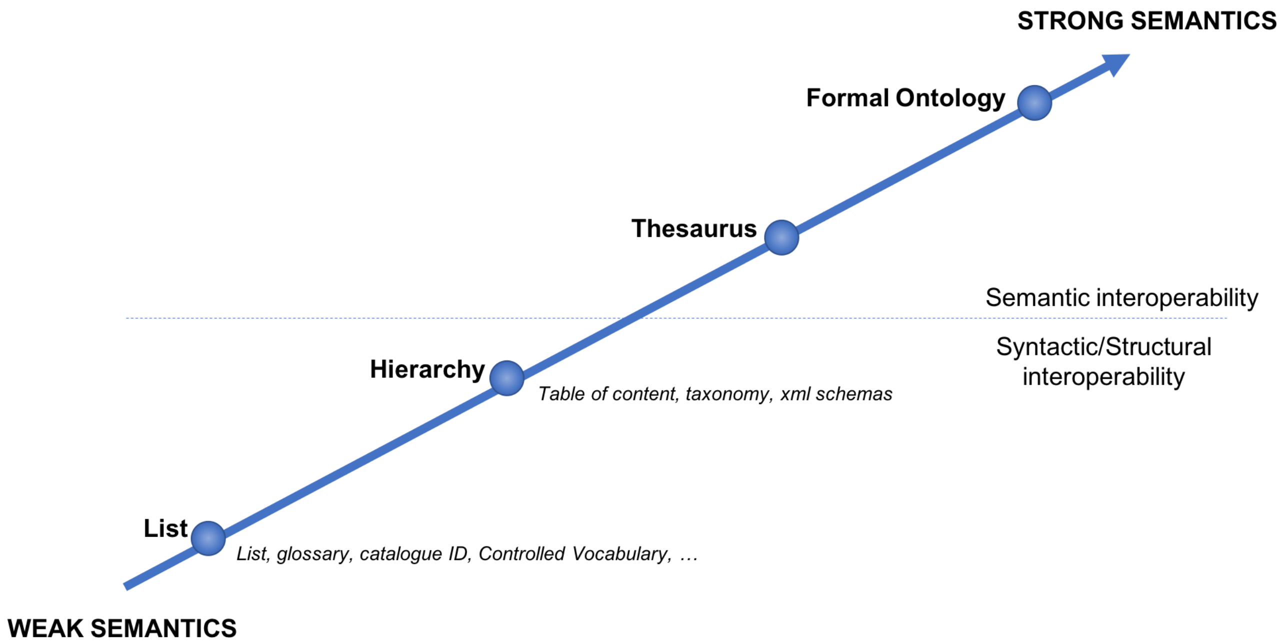


*Figure 2: Semantic artefact spectrum. Derived from Leo Obrst, 2010*

*Semantic artefacts are classified into 4 main types: list, hierarchy, thesaurus and formal ontology. These 4 different types of semantic artefacts are represented along an axis going from "weak semantics" to "strong semantics". Examples of subtypes are provided on the right side of the axis. A dichotomy can be made between hierarchy and thesaurus. On the one side the simplest types are supporting syntactic interoperability allowing machines to process information due to compatible syntax. On the other side, semantic interoperability is being achieved allowing machines to interpret and reason over the data.*

The second block corresponding to hierarchical models (informal hierarchies and taxonomies) builds upon a list of terms/concepts organised hierarchically using either "loose" parent/child or the more formal "*is a*" relations. These hierarchies can then be enriched with additional relations such as synonyms and association relations therefore becoming thesauri. Thesauri can be used as a basis to create formal ontologies by adding axioms and rules. This type of "Russian doll" like organisation is shown in fig. 3. It allows us to visualise a path of transformation between semantic artefact types.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

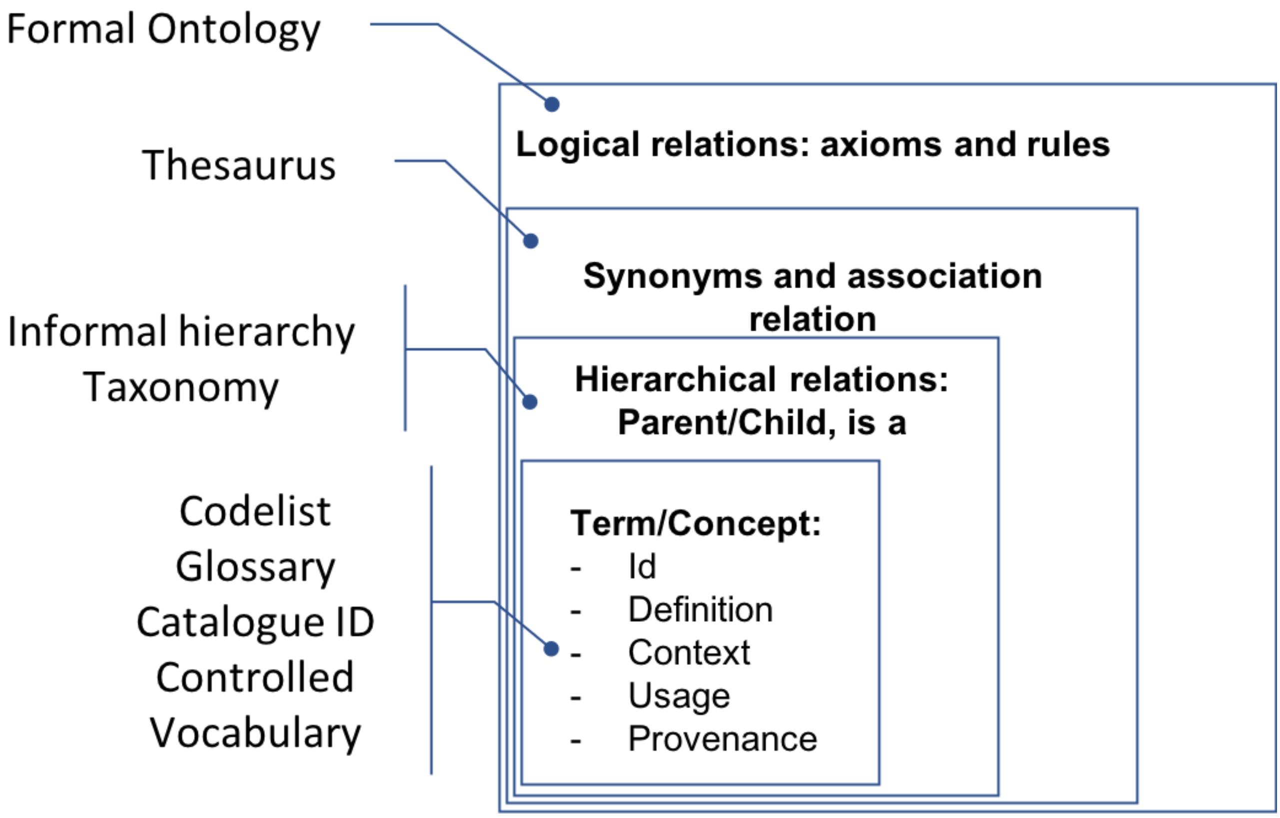


*Figure 3: From list to formal ontology: a transformation path.*

In addition to the complexity, semantic artefacts are also heterogeneous in nature due to the diversity of the data models and standards used to serialise semantic models. Various standards for data models (RDF[14], RDFs[15], OWL[16], SKOS[17]) and serialisation formats ( XML[18], XML Schema[19], JSON, RDF/XML[20], OWL/XML[21], Manchester Syntax[22], JSON-LD[23], Turtle[24], N-Triples[25]) have been proposed by W3C . For simple models, common formats such as XML, XML Schema and JSON are typically used. As the model becomes more complex, more expressive data models have been proposed. The Resource Description Framework (RDF) is one of them. It is a formal language for describing

[14] RDF https://www.w3.org/TR/rdf11-concepts/
[15] RDFs https://www.w3.org/TR/2014/REC-rdf-schema-20140225/
[16] OWL https://www.w3.org/OWL/
[17] SKOS https://www.w3.org/TR/skos-primer/
[18] XML https://www.w3.org/TR/xml/
[19] XML Schema https://www.w3.org/TR/xmlschema-0/
[20] RDF/XML https://www.w3.org/TR/rdf-syntax-grammar/
[21] OWL/XML https://www.w3.org/TR/2012/REC-owl2-xml-serialization-20121211/
[22] OWL Manchester Syntax https://www.w3.org/TR/2012/NOTE-owl2-manchester-syntax-20121211/
[23] JSON-LD https://www.w3.org/TR/json-ld11/
[24] Turtle https://www.w3.org/TR/turtle/
[25] N-Triples https://www.w3.org/TR/2014/REC-n-triples-20140225/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

information as a very simple graph-oriented data schema. Based on a URI to identify the resources, RDF enables the exchange of data on the Web between applications, while preserving their original meaning and facilitating the processing and re-combination of the contained information. There are many serialisations of RDF such as RDF/XML, Turtle, JSON-LD, N3, etc. RDF is complemented by RDF Schema denoted as RDFs. This extra layer provides additional data-modelling elements for RDF data thus extending the expressivity of the supported models.

To support a higher level of expressivity and logic to represent complex semantic models, W3C proposed the Web Ontology Language (OWL). This language extends the couple RDF/RDFs with additional reasoning options grounded in formal logic. OWL exists in various flavours and expressivity profiles (e.g. OWL-Lite, OWL-full, OWL-DL). Despite these powerful additions enabling reasoning and automated processing, OWL suffers a limitation due to the initial working hypothesis used to formalise the logic. Indeed based on the Open World Assumption, OWL cannot represent closed logic. To address this issue, a recent addition to the W3C standards has been proposed, SHACL (Shapes Constraint Language)[26]. This Language allows to specify constraints on RDF graphs.

Finally another standard used to build semantic artefacts is the Simple Knowledge Organisation System (SKOS)[27]. This standard, less formal and constrained than RDF and OWL, is also a W3C recommendation to build Knowledge Organisation Systems (KOS) (i.e. semantic artefacts). SKOS is quite popular for building thesauri, classification schemes, subject heading systems and taxonomies (as shown in Table 1). The main reason for such popularity lies in the fact that it has no formal grounding and people use it to express all kinds of containment relations. For example the skos: broaderProperty is used to express a subclass relation (mammal skos:broader animal), a subregion (Texas skos:broader USA), subperiod (baby-boom-period skos:broader 20thCentury) etc. Despite the lack of formal grounding, most humans do understand the inherent reasoning and can develop in retrospect applications that properly deal with these mappings. SKOS provides a standard way to represent knowledge organisation systems using RDF, allowing them to be passed between computer applications in an interoperable way and to be used in distributed, decentralised metadata applications, where metadata is harvested from multiple sources.

In table 1 below, we are listing the common formats/standards used to build each of the 4 types of semantic artefacts.

| Type of Semantic artefact | Currently used standards (serialisation formats and data models) |
|---|---|
| List (terminologies, glossaries, vocabularies) | CSV, XML, JSON, SKOS |
| Hierarchical list | XML schema, RDF schema, SKOS |

[26] https://www.w3.org/TR/shacl/

[27] SKOS, Simple Knowledge Organisation System, https://www.w3.org/2004/02/skos/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| | |
|---|---|
| Thesaurus | RDF/RDFs, SKOS |
| Formal ontology | OWL, OntoUML, FOL, Modal logic |

*Table 1: Association between the type of semantic artefact and standards used.*

## 1.2. What is meant by FAIR Semantics?

FAIR Semantics, in the context of this project, means that semantic artefacts should adhere to the FAIR principles. For this, we are considering semantic artefacts as a specific type of data, used to describe or annotate other data, i.e. as metadata. This also reflects a transition from the common understanding of metadata based on controlled vocabularies, tags, and labels, towards 'next generation metadata' expressed as Linked Data or Permanent Identifiers (Smith-Yoshimura, 2020). This approach allows us to consider each individual FAIR principle in the context of semantic artefacts. This implies the following:

- usage of globally unique persistent and resolvable identifiers for semantic artefacts, their content (i.e. concept/term/class and relation) and their version,
- machine-readable metadata to describe the semantic artefacts themselves and their content,
- usage of repositories to share, publish and retrieve semantic artefacts and their content
- defining common API(s) to access and index semantic artefacts and their content,
- interoperability approaches to make sure that semantic artefacts of various degrees of complexity and encoding format should work together including publishing mappings and crosswalks between semantic artefacts,
- semantic artefacts and their content should be retrievable through search engines.

Solutions to address part or all these issues have been developed within domain specific communities. As our goal is not to reinvent but rather reuse, we are providing together with our recommendation pointers to existing community-specific recommendations. As of now, we have included the following recommendations:

- OBO Foundry (Smith et al., 2007)
- Industry Ontology Foundry (Kulvatunyou et al., 2018)
- Agrisemantic Working Group recommendations: 39 Hints to Facilitate the Use of Semantics for Data on Agriculture and Nutrition[28].
- Metadata for Ontology Description and Publication Ontology [MOD] (Dutta et al., 2017)
- Ontology Metadata Vocabulary [OMV] (Hartmann et al., 2005)

[28] Agrisemantic Working Group https://agrisemantics.org/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

- Minimum Information for Reporting an Ontology [MIRO] (Matentzoglu et al., 2018)
- Minimum Information to Reference an External Ontology Term [MIREOT] (Courtot et al., 2009)
- Linked Open Vocabulary [LOV] (Vandenbussche et al., 2017)
- Best practices for implementing FAIR vocabularies and ontologies on the Web. (Garijo & Poveda-Villalón, 2020)

| |
|---|
| Community experts are invited to contribute to extend our view of community practices by adding any missing recommendation. Please do so by adding requests and suggestions in github:<br>https://github.com/FAIRsFAIR/FAIRSemantics |

## 1.3. Stakeholder Groups

The goal of this deliverable, and the FAIR semantics task, is to co-create both recommendations for making semantic artefacts FAIR and a set of agreed best practices to follow together with the community of semantics at large. While working on this initial set of recommendations, we realised that recommendations are often targeted to particular stakeholders. Some recommendations are very specific about the format, structure and content of semantic artefacts, therefore useful for practitioners. Some others are also directed towards developers and maintainers of semantic repositories, while some recommendations actually highlight the need for a community wide consensus to fill gaps in the current landscape of standards and data models for semantic interoperability.

Therefore, for this phase of the work, we are considering three main stakeholder groups:

1. Expert vocabularies/ontologies managers, practitioners dealing with the creation and maintenance of the semantic artefacts that are called in this document Semantic Artefact developers/managers
2. Repositories managers i.e. development team and curators of community specific semantic repositories;
3. Semantic Web Community at large, dealing with semantic artefacts in general, in different contexts, including research data infrastructures, etc.

For each of the recommendations, we listed the impacted/ concerned stakeholder groups and we are providing summary tables for each of stakeholders, listing the recommendations of interest.



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

These recommendations are preliminary. We are inviting all the interested stakeholders to join our effort and contribute to establish common guidelines for harmonising the semantic landscape.

### 1.4 What changed?

Since the publication of the second version, we received a lot of relevant inputs and feedback about our work. Unfortunately, we were not able to address them all. In this version of the document, we have made minor updates of the recommendations and removed three best practices which were not clear. In addition, the recommendations are grouped along 4 main topics: identifier, metadata, repository and semantic interoperability. The fact that we only made minor updates on the individual recommendations makes us believe that we reached a first consensus which should be benchmarked against the reality. Some initiatives already leveraged these recommendations in their work. We added a short analysis of the current uptake in section 5. We also analysed the existing best practices and refined our best practices to align with the existing ones.

One of the main inputs we received was about the practicalities of the recommendations. How to implement them? As this is clearly a daunting task, we focused on 2 specific aspects. First, the definition of a generic service architecture to leverage FAIR Semantics based on a driving use-case and second, we attempted to resolve the issue with respect to the metadata. Indeed, FAIR principles mention that a rich metadata description should be provided. What is meant by rich metadata for Semantic artefact? Can we define a minimum metadata schema which will act as a threshold to FAIRness? Below this threshold, semantic artefacts are not FAIR and above they are more FAIR. These two additions are described in section 4.

## 2. Recommendations

In this section, we are providing the final version of the list of individual recommendations. The recommendations were originally derived from the material we gathered during the workshop that took place as a co-located event to the RDA 14 Plenary in Espoo on 23 October 2019. In the workshop more than 20 experts brainstormed and discussed the different criteria of FAIR, elaborating on the implications of these requirements on semantic artefacts. From this material the FAIRsFAIR task group formulated more than 40 recommendations/requests/requirements, which were then analysed individually. We evaluated how such input relates to one or more particular FAIR principles and aggregated them whenever possible into one recommendation. The result was published as D2.2: 1st set of Recommendations for FAIR Semantics (Le Franc et al., 2020). Feedback has then been collected in discussions in different expert fora, such as the RDA Plenaries 15 and 16, as well as through the fairsfair.eu web page for community review, the RDA VSSIG 'FAIR Semantic Repositories Task Group' and on github. The most extensive feedback was collected in a workshop on October 15th 2020. Based on this feedback, the 17 preliminary recommendations presented in the previous iteration have been slightly reviewed, completed and updated. The recommendations are each enriched with a description providing some context, and whenever possible, existing



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

recommendations. In addition, we have been considering the FAIR principle(s) addressed by recommendations, and we also consider which stakeholders are impacted/ responsible for such recommendations.

The recommendations that could not be directly aligned to the FAIR principles were aggregated into a set of suggested best practices presented in the next section. These Best Practice Recommendations are not directly linked to a particular recommendation but contribute to improve the overall FAIRness of semantic artefacts.

Individual recommendations are represented as tables in which the top row contains the number of the recommendation, the recommendation and the associated FAIR principles in a dedicated cell. The other rows contains as follows:

1. Description
2. Existing Recommendations/ Existing technologies
3. Stakeholder
4. Examples

In this final version of the recommendations, we are presenting them organised along 4 main topics: Globally Unique Persistent and Resolvable Identifier, Metadata, Repository and Semantic Interoperability.



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## 2.1 Globally Unique Persistent and Resolvable Identifier

| | | |
|---|---|---|
| Mandatory | P-Rec. 1: Globally Unique, Persistent and Resolvable Identifiers must be used for Semantic Artefacts, their content (terms/ concepts/ classes and relations), and their versions. | F1 |

**Description**

Semantic artefacts are typically structured text files. They are *de facto* digital objects and should be unambiguously identified by Globally Unique, Persistent and Resolvable Identifiers (GUPRI) as specified in the EOSC PID policy recommendations (Hellström et al., 2020) . In the context of a web of FAIR data, these identifiers should be resolvable and support the retrieval of both the semantic artefact itself and also its metadata (see Rec. 2 regarding metadata). As shown in fig. 1, semantic artefacts are composite digital objects requiring at least three levels of identifiers: one for the semantic artefact itself, one for its content and one for the metadata (including both the global metadata and the metadata associated with the content). The latter is described in the following recommendation (Rec. 2). Finally, semantic artefacts are living digital objects by nature, evolving over time. Each version of a semantic artefacts should be uniquely identifiable, allowing access to the latest version by default but also providing access to previous versions in use in existing information systems.

As discussed in the introduction, these identifiers should apply to the semantic artefact but also to its content. Indeed, semantic artefacts can be considered as collections of concepts and relations (datasets). Therefore, in this context, each element/ component of the semantic artefact should also have an associated GUPRI i.e. each term/ concept/ class

Finally, a unified identifier schema should be used to identify each version of semantic artefact. This can be done using versioned URI as proposed by OBO Foundry. Using GUPRI for the different version allows information systems to automatically retrieve the latest version and older versions of the semantic artefact.

This recommendation emphasises the need for reliable and persistent identification systems without any technical constraints.

**Existing Recommendations**

- W3C Data on the Web - Best Practice 9: Use persistent URIs as identifiers of datasets namespace[29]
- OBO Foundry - Principle 3[30]
- OBO Foundry - Identifier Policy[31]
- OBO Foundry - Principle 4[32]
- Industrial Ontology Foundry - principle 11 IRI and identifier space
- Industrial Ontology Foundry - principle 12 Identifier and naming conventions
- EOSC PID policy recommendation (Hellström et al., 2020)

[29] W3C Data on the Web - Best Practice 9 https://www.w3.org/TR/dwbp/#DataIdentifiers
[30] OBO Foundry principle 3 http://www.obofoundry.org/principles/fp-003-uris.html
[31] OBO Foundry ID policy http://www.obofoundry.org/id-policy
[32] OBO Foundry principle 4 http://www.obofoundry.org/principles/fp-004-versioning.html



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| • Garijo & Poveda; Best practices for Implementing FAIR Vocabularies and Ontologies on the Web Principles 2., 2.1, 2.2 and 2.3 |
| **Stakeholders** |
| Practitioner, Repository |

| | | |
|---|---|---|
| Mandatory | P-Rec. 2: Globally Unique, Persistent and Resolvable Identifiers must be used for Semantic Artefact Metadata Records. Metadata and data must be published separately, even if it is managed jointly. | F1, F3 |
| **Description** | | |
| Semantic artefacts are often built as containers including both their descriptive metadata and data. Commonly, semantic artefacts contain a set of concepts and their descriptions and are identified by a URL pointing to a file to download which should be parsed to access the content including the metadata. These practices contribute to the lack of findability by hiding the metadata to machines. For this purpose, it is necessary to consider publishing the ontology metadata separately allowing potential users to find it.<br><br>This metadata record should also have a GUPRI (Globally Unique, Persistent and Resolvable Identifier - FAIR principle F1) and an explicit reference to the semantic artefact it describes (FAIR principle F3). In this way, search engines can retrieve and index metadata that uniquely point to their related semantic artefacts. This recommendation puts an emphasis on the necessity to publish metadata separately from the semantic artefact and have services to share/ publish ontologies which should support the extraction and the publication of their metadata as suggested in P-Rec. 4. | | |
| **Stakeholders** | | |
| Practitioner, Repository | | |

## 2.2 Metadata

| | | |
|---|---|---|
| Mandatory | P-Rec. 3: A common minimum metadata schema must be used to describe semantic artefacts and their content | F2, R1.1, R1.2 and R1.3 |
| **Description** | | |
| As with any type of data, semantic artefacts should be described by different levels of metadata to allow users to retrieve them and to understand their content. In particular, it is important to have general information regarding the scope of the semantic artefact (at least which domain is covered by the ontology), provenance information and many other details. Metadata must be appropriate to the life cycle stage and application of the artefact. | | |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

This metadata should be available in popular encodings and should be accessible for harvesting and discovery by search engines, semantic service providers and metadata aggregators, registries, and catalogues.

Unfortunately, there is currently no consensus on a common set of metadata elements to describe semantic artefacts. Several initiatives are proposing their recommendations such as OBO Foundry and IOF and several metadata schemata have been developed such as LOV (Vandenbussche et al., 2017), Ontology Metadata Vocabulary (OMV)[33], Metadata for Ontology Description and Publication Ontology (MOD),... (see list of related recommendations below). However, the heterogeneity of these metadata schema hampers indexing, and retrieval, as well as reuse of the semantic artefacts.

To address this issue, we defined, together with the community, a preliminary minimum set of mandatory properties defined within a DCAT application profile for describing Semantic Artefacts[34]. The process and the model are defined in section 4.2.

As for semantic artefact themselves, the concept /term/ class and relation that compose them should also have a common metadata schema that provide information such as label, definition, examples of usage, author, version, multilingual labels, and similar.

Reaching an agreement at this level will ease the process of working with concepts from multiple heterogeneous semantic artefacts. It is important to note that proper definitions are necessary to be able to evaluate the difference between similar classes from different ontologies (see BP-Rec. 8).

**Existing Recommendations**

- OBO Foundry - Principle 8 Documentation[35]
- OBO Foundry - Principle 5 Scope[36]
- OBO Foundry - Principle 6 Textual definition[37]
- Industry Ontology Foundry - Requirement 9 Documentation[38]
- Industry Ontology Foundry - Requirement 5 Scope[39]
- LOV - DCAT based metadata schema
- VOAF[40]
- Ontology Metadata Vocabulary[41]
- Metadata for Ontology Description and Publication Ontology[42]

[33] OMV http://mayor2.dia.fi.upm.es/oeg-upm/index.php/en/downloads/75-omv/index.html
[34] https://github.com/FAIRsFAIR/SemanticDCAT-AP
[35] OBO Foundry principle 8 http://www.obofoundry.org/principles/fp-008-documented.html
[36] OBO Foundry principle 5 http://www.obofoundry.org/principles/fp-005-delineated-content.html
[37] OBO Foundry principle 6 http://www.obofoundry.org/principles/fp-006-textual-definitions.html
[38] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[39] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[40] VOAF https://lov.linkeddata.es/vocommons/voaf/
[41] Ontology Metadata Vocabulary http://omv2.sourceforge.net/
[42] MOD-Ontology https://github.com/sifrproject/MOD-Ontology



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| • W3C Data on the web best practices - BP1, BP2 and BP3[43]<br>• Networked Knowledge Organisation Systems Dublin Core Application Profile (NKOS AP)[44] |
| **Stakeholders** |
| Practitioner, Repository and Community |

| Mandatory | P-Rec. 8: Human and machine-readable persistence policies for semantic artefacts metadata and data must be defined. | A2 |
|---|---|---|
| **Description** | | |
| Once published in a semantic repository, semantic artefacts will be reused by others to build their information systems. In the eventuality where the semantic artefact, a concept/ term or a relation is deprecated or simply replaced, the repository should offer persistence policies for the metadata (duration of archiving, …). These policies should be both human and machine readable. Machine readable policies will allow services to automatically detect the change, to either warn the user or to directly integrate the change whenever it is possible. For humans, repositories could use the classical tombstone page with redirects to the new page when the semantic artefact or the element has been replaced.<br><br>Semantic Artefact components (terms/ concepts/ classes) with compliant metadata and GUPRI structures (P-Rec 1, 2, 3) will not, under typical circumstances, be deprecated, but replaced by new versions. If community recommendations similar to those for data are followed for version management, these objects will remain available (Rauber et al., 2015). | | |
| **Existing Recommendations** | | |
| RDA - Recommendations on Citation of Evolving Data (Rauber et al., 2015) | | |
| **Stakeholder** | | |
| Repository, Community, Practitioner | | |

| Mandatory | P-Rec. 9: Semantic artefacts must be made available as a minimum portfolio of common serialisation formats. | I1 |
|---|---|---|
| **Description** | | |
| Semantic artefacts should be serialised in one or more of the formats developed in the context of the Semantic Web and Linked Data i.e. OWL, OBO, RDF, and SKOS. Semantic repositories should provide a Linked Data compliant API to enable the creation of a semantic graph for analysis and reuse.<br><br>However, these standardised formats have limited capabilities for serialising highly expressive logical | | |

[43] Data on the Web Best Practices https://www.w3.org/TR/dwbp/
[44] NKOS AP https://nkos.slis.kent.edu/nkos-ap.html



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| models . A good practice should be to share a serialisation of highly expressive model to provide at least access to the concepts/terms (e.g. OWL-Full, OWL-DL,..). However, this implies a loss of information and expressivity. In P-Rec. 11, we are recommending the community to define a common standard and a serialisation format to describe complex logical relations. |
| **Existing Recommendations** |
| • RDF and RDFS[45]<br>• W3C-OWL (OWL 2 / Full; OWL 2 / EL; OWL 2 / QL; OWL 2 / RL) stack[46]<br>• SKOS[47]<br>• The OBO Foundry Principle 2 "Common Format" requires the OWL file in RDF-XML format. Legacy formats are automatically converted to OWL.[48]<br>• Industry Ontology Foundry - requirement 3[49]<br>• OBO<br>• Linked Data Platform |
| **Stakeholder** |
| Practitioner, Repository |

| Optional | P-Rec. 14: Standard vocabularies should be used to describe semantic artefacts | I2 |
|---|---|---|
| **Description** | | |
| As stated in Rec. 3, the semantic artefact metadata is important for both findability and reusability. Standard vocabularies used by metadata schema, such as DCAT, Dublin Core and FOAF for example, should be used to describe such semantic artefacts. Agreeing on a common set of standard vocabularies would allow to improve the interoperability of the metadata descriptions. This should also apply to the metadata associated with the content.<br><br>In section 4.2, we are describing the FAIRsFAIR DCAT Application Profile for FAIR Semantic artefacts which leverages a set of standard vocabularies for describing semantic artefacts. This first version has been defined together with the community of experts and is open to comments. | | |
| **Existing Recommendations** | | |
| • LOV recommendations[50]<br>• FDP metadata scheme[51] | | |

[45] RDF and RDFS https://www.w3.org/TR/rdf11-primer/
[46] W3C-OWL https://www.w3.org/TR/owl2-profiles/
[47] SKOS https://www.w3.org/TR/skos-primer/
[48] The OBO Foundry principle 2 http://www.obofoundry.org/principles/fp-002-format.html
[49] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[50] LOV https://lov.linkeddata.es/dataset/lov/
[51] FAIR Data point specification https://github.com/FAIRDataTeam/FAIRDataPoint/wiki/FAIR-Data-Point-Specification



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| ● OBO Foundry<br>● FAIRsFAIR SemanticDCAT-AP |
| **Stakeholder** |
| Practitioner, Community |

| Optional | P-Rec. 15: Provenance information regarding the reuse of components from third-party semantic artefacts should be made explicit | I3, R1.2 |
|---|---|---|
| **Description** | | |
| New semantic artefacts can be built upon existing artefacts. In some cases reuse involves copying and pasting elements of one artefact to another. This mechanism does not allow automatic access to reused elements and their semantic artefact. In order to be able track the element reused from other semantic artefacts, reference to third party semantic artefacts should be made explicit. For this, external semantic artefacts (in part or in full) should be imported using a specific metadata element to represent the import (e.g. <owl:import/>). When using such explicit references, it becomes possible to extract this dependency information automatically from the artefacts.<br><br>The explicit reference can be made<br>● by providing a direct link to the used resource, preferably as a GUPRI<br>● or by describing the link as inbound and outbound links as proposed by LOV<br>● or by using VoID vocabulary to interlink the different ontologies<br>● or by considering the requirement of the Minimum Information to Reference an External OnTology, MIREOT (Courtot et al., 2009).<br><br>This recommendation does not pretend to enforce a particular solution but rather emphasises the need for the community of semantic web to define a common way of referencing or importing external semantic artefacts. | | |
| **Existing Recommendations** | | |
| ● LOV[52]<br>● VoID[53]<br>● MIREOT[54] | | |
| **Stakeholder** | | |
| Practitioner | | |

[52] LOV https://lov.linkeddata.es/dataset/lov/
[53] VoID https://www.w3.org/TR/void/
[54] MIREOT http://precedings.nature.com/documents/3574/version/1



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Mandatory | P-Rec. 16: The semantic artefact must be clearly licenced for use by machines and humans | R1.1 |
|---|---|---|
| **Description** | | |
| Proper reuse of digital objects requires a human and machine-readable licence. Well documented legal interoperability is a prerequisite for automatic distributed search and the use of both the semantic artefacts and their component terms/concepts/classes and relations.<br><br>Although we are encouraging Open licences, preferably using Creative Commons 4.0 licensing, this recommendation doesn't impact the choice but emphasises the need for adding this information explicitly for both human and machine to avoid ambiguities on the conditions for reuse. See BP-Rec. 12 for more information. | | |
| **Relevant framework** | | |
| ● Creative Commons licences[55]<br>● ODRL[56] | | |
| **Stakeholder** | | |
| Practitioner, Repository | | |

| Mandatory | P-Rec. 17: Provenance must be clear for both humans and machines | R1.2 |
|---|---|---|
| **Description** | | |
| Semantic artefacts are living digital entities undergoing changes and revisions to cope with semantic drift and for improving/ extending the scope or granularity. Provenance information describing all these changes during the semantic artefact lifecycle should be provided to potential external users. This information can be thus used to evaluate the semantic artefact and understand the release cycle.<br><br>Provenance should be documented at an appropriate level of granularity to enable reuse of semantic artefacts and its constituting elements (class/ term and relation). Provenance should be presented to the human user but also should be expressed in a machine-readable way. All appropriate sources should be referred to (both source PID (data object) and creator PID) and the provenance should provide dates and lifecycle events.<br><br>Provenance information should be described using an appropriate standard model such as PROV[57]. PROV-based machine readable description could be then used to provide means to automatically update any resource using the semantic artefact. The provenance information should contain all the necessary elements to build representations to the users such as changelogs and describe backward interoperability. | | |

[55] Creative Commons licences https://creativecommons.org in rdf https://github.com/creativecommons/cc.licenserdf
[56] ODRL https://www.w3.org/TR/odrl-model/
[57] PROV data model https://www.w3.org/TR/prov-primer/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Existing Recommendations |
|---|
| - MIRO (Matentzoglu et al., 2018)<br>- OBO foundry - Principle 4[58]<br>- OBO Foundry - Principle 8[59]<br>- PROV[60]<br>- Metadata schema elements (D-CAT, DataCite) |
| **Stakeholder** |
| Practitioner, Repository |

## 2.3 Semantic Repositories

| Optional | P-Rec. 4: Semantic Artefact and its content should be published in an appropriate semantic repository | F4 |
|---|---|---|
| **Description** | | |
| Semantic artefacts are made accessible using a wide variety of mechanisms, including publication in open repositories such as Zenodo or Figshare, deployment to github, availability as a downloadable file or object in a website, or a response offered by a web service. Most of the time semantic artefacts need to be downloaded and parsed in order to have access to its content, such as concepts/ terms, relations, and metadata.<br><br>This hampers the findability of the semantic artefact and makes reuse more difficult (see Rec. 2). To solve these issues, specific repository technologies have been developed to support the publication of semantic artefacts, their content and the metadata associated with the semantic artefacts. These "semantic repositories" provide interfaces for both humans and machines to consume semantic artefacts. They are an important piece of the infrastructure underlying the implementation of FAIR principles and FAIR Semantics as pointed out in the "Turning FAIR into reality" report and action plan (European Commission Expert Group on FAIR Data, 2018).<br><br>The number of such repositories is currently increasing with domain specific repositories and registries such as Bioportal, EBI-OLS, Ecoportal[61], Agroportal[62], BODC NERC vocabulary service[63] or more generic services such as Finto.fi[64], BARTOC[65] or Research Vocabularies Australia[66]. | | |

[58] OBO foundry principle 4 http://www.obofoundry.org/principles/fp-004-versioning.html
[59] OBO foundry principle 8 http://www.obofoundry.org/principles/fp-008-documented.html
[60] https://www.w3.org/TR/prov-primer/
[61] Ecoportal http://ecoportal.lifewatchitaly.eu/
[62] AgroPortal http://agroportal.lirmm.fr/
[63] BODC vocabulary service http://seadatanet.maris2.nl/v_bodc_vocab_v2/welcome.asp
[64] finto.fi http://finto.fi/en/
[65] BARTOC https://bartoc.org
[66] Research Vocabularies Australia https://vocabs.ands.org.au/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| These repositories should act as a trustworthy long term archive (ideally certified by e.g. CoreTrustSeal), should provide GUPRIs, publish metadata making the semantic artefact findable for humans through a dedicated User Interface, and for machines through an API. Trustworthy data repositories, focused on specific disciplines, are available for data curation and preservation, and these repositories may be amenable to preservation of semantic artefacts in the short term. In the medium term, development of criteria for Trustworthy Semantic Repositories is a community responsibility, See BP-Rec. 11.<br><br>This recommendation does not aim to support any particular technology but emphasises the necessity to share/publish and preserve semantic artefacts in such repositories to improve both findability and reuse over time. |
| **Existing technologies** |
| ● SKOSMOS[67]<br>● Bioportal[68]<br>● Research Vocabularies Australia |
| **Stakeholders** |
| Practitioner, Community |

| Mandatory | P-Rec.5: Semantic repositories must offer a common API to access Semantic Artefacts and their content in various serialisations for both use/ reuse and indexation by any search engines | F4, A1, A1.1 |
|---|---|---|
| **Description** | | |
| Semantic artefacts are distributed across the web in a variety of locations and formats. Semantic repositories act as aggregators of semantic artefact publishing both the metadata and the content of the semantic artefacts and providing a search engine and an API to search and access the content through dedicated services. However, the APIs to search and access content is specific to each repository which hampers the possibility to access content from multiple sources for use and reuse but also for indexing by search engine. Part of the API heterogeneity is linked with the diverse metadata schema used by repositories to describe semantic artefacts.<br><br>To enable federated searches across repositories, it is necessary to harmonise the API landscape by defining a common set of API features based on a common minimum set of metadata for describing semantic artefacts (see P-Rec. 3 and section 4.2) and their distribution (P-Rec.9).<br><br>Enabling automated indexing across repositories will require agents to access a machine readable description of the API and the description of information that can be accessed. Therefore, repositories | | |

[67] SKOSMOS https://www.kansalliskirjasto.fi/en/services/system-platform-services/skosmos

[68] Bioportal https://bioportal.bioontology.org/;



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| should consider publishing at least the description of their API using OpenAPI specifications which will provide first a human readable API documentation in a machine readable format. A recent extension of the OpenAPI specification, called smartAPI[69] has been proposed to provide semantically annotated API description to make an API FAIR. Such semantically enriched description could enable automated workflows for indexing semantic repositories.<br><br>Several other possible solutions exist i.e. publishing directly the content as Linked Data and to be compliant with the LD standards (Linked Data Platform, …), use a common inter-exchange metadata format such as RIF-CS or publish metadata and content using the FAIR Data Point service.<br><br>This recommendation does not aim to support any particular solution but rather emphasise the need for the community of semantic repositories to agree on a common solution. |
| **Existing Recommendations** |
| • The FAIR Data point specification[70]<br>• Registry Interchange Format - Collections and Services (RIF-CS) (ISO 2146)[71].<br>• Linked Data Platform[72]<br>• OpenAPI[73]<br>• smartAPI[74] |
| **Stakeholder** |
| Repository, Community |

| Optional | P-Rec. 6: Build semantic artefact search engines that operate across different semantic repositories | F4 |
|---|---|---|
| **Description** | | |
| To be able to reuse existing semantic artefacts in part or in full, it is necessary to be able to find them across a large number of distributed and heterogeneous semantic repositories. For this, we need semantic artefact search engines such as the discontinued Swoogle[75] that can operate across different semantic repositories. These search engines will enable federated queries across the semantic artefacts and provide means to gather analytics across all ontologies (overlaps, mappings, reuse) and large scale automated mappings to resolve semantic ambiguity. The indices supporting the search engines could also be directly integrated within the tooling to provide access to the existing resource at the time of the creation of a new semantic artefact, or be used to populate lookup lists and vocabulary resources for applications. . | | |

[69] smartAPI https://smart-api.info/
[70] FAIR Data point specification https://github.com/FAIRDataTeam/FAIRDataPoint-Spec
[71] RIF-CS https://vocabs.ands.org.au/viewById/1
[72] Linked Data Platform https://www.w3.org/TR/ldp/
[73] OpenAPI https://www.openapis.org/
[74] smartAPI https://smart-api.info/
[75] https://ebiquity.umbc.edu/project/html/id/53/Swoogle



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| This recommendation emphasises that such service is an important element of the infrastructure to support FAIR data and FAIR Semantics. |
| **Existing Recommendations** |
| N/A |
| **Stakeholder** |
| Community, Repositories |

| Mandatory | P-Rec. 7: Repositories should offer a secure protocol and user access control functionalities | A1.2 |
|---|---|---|
| **Description** | | |
| Semantic artefacts should be openly shared to support reuse and to avoid concept redundancy and semantic ambiguities.<br><br>However, semantic artefacts might be developed under specific copyrights with paywalls (e.g. Dewey Decimal Classification) preventing direct access for use. Although the metadata describing the artefact should be made available, the access can be restricted with user access control functionalities and secured access to the content should be provided through a secure protocol such as HTTPS for machines.<br><br>These access condition should be explicitly defined in the semantic artefact metadata as proposed in the DCAT-AP for Semantic Artefact (see section 4.2) | | |
| **Existing Recommendations** | | |
| N/A | | |
| **Stakeholder** | | |
| Repository | | |

## 2.4 Semantic Interoperability

| Optional | P-Rec. 10: Foundational Ontologies may be used to align semantic artefacts | I1, I2, I3 |
|---|---|---|
| **Description** | | |
| Foundational ontologies are complex logical representations of the basic concepts of the world of discourse. Grounding domain-specific semantic artefacts in foundational ontologies allows the alignment of | | |



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| various domain specific semantic artefacts around a common hypothesis about the world. These semantic artefacts are built to support the integration and the interoperation[76] of domain specific semantic artefacts acting as a language bridging them. Several foundational ontologies exist, such as UFO[77], BFO[78], DOLCE[79], EMMO[80].<br><br>This recommendation does not make any claim regarding which foundational ontology to use but emphasises the value of being aligned with one.<br><br>This recommendation is more appropriate for formal semantic artefacts (ontologies, thesaurus, taxonomies) where such alignment is strongly recommended. Vocabulary and linked lists may be expedient and the recommendation is less critical in such cases. |
| **Existing Recommendations** |
| • Industry Ontology Foundry - requirement 8[81] |
| **Stakeholder** |
| Practitioners |

| Optional | P-Rec 11: A standardised language should be used for describing high expressivity semantic artefacts. | I1 |
|---|---|---|
| **Description** | | |
| As discussed in P-Rec. 9, knowledge representation languages such as OWL and RDF, commonly used to represent semantic artefacts by the Semantic Web and Linked Data communities, cannot express all characteristics of more complex/ expressive semantic models. This lack of standard coverage impedes interoperability as complex semantic artefacts have to be simplified, resulting in a loss of information and expressivity. To address this issue, the semantic web community should define an additional common language for high expressivity model representation as per FAIR principle I1. | | |
| **Existing Recommendations** | | |
| • SHACL[82]<br>• SWRL[83] | | |

[76] In other words: Semantic Interoperability is the main benefit from this recommendation
[77] UFO https://ontouml.readthedocs.io/en/latest/intro/ufo.html
[78] BFO https://basic-formal-ontology.org/
[79] DOLCE http://www.loa.istc.cnr.it/dolce/overview.html
[80] EMMO https://emmc.info/taxonda/emmo-european-materials-modelling-ontology/
[81] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[82] SHACL https://www.w3.org/TR/shacl/
[83] SWRL https://www.w3.org/Submission/SWRL/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| • OntoUML[84] |
| **Stakeholder** |
| Community |

| Optional | P-Rec. 12: Semantic mappings between the different elements of semantic artefacts should be serialised in machine-readable formats | I1, I3, R1.3 |
|---|---|---|
| **Description** | | |
| As we discussed in P-Rec. 10, semantic artefacts are often developed to describe a specific aspect of a scientific domain. Despite this reduced scope, several models of the same aspects can co-exist. They are either developed *de novo* or developed as a part of another ontology. This duplication is often due to a lack of knowledge regarding existing semantic artefacts. In order to aggregate distributed resources aligned with these different models, it is necessary to create mapping relations between the elements of such semantic artefacts.<br><br>In many cases, these mappings are based on existing relations (such as *sameAs* from SKOS). However, mappings can become complex especially when considering logical relations. Risks include representing content drift, as well as context insensitive use of semantic artefacts, and there are no common descriptions for such complex mappings. Mappings are often created by individuals for satisfying a specific need. Information regarding the provenance and usage of these mappings are of importance for any practitioner who would be interested in reusing them.<br><br>This recommendation aims at highlighting a gap in relation to the achievement of principle I3 and to emphasise the need for harmonisation and the implementation of machine readable descriptions of mapping in order to foster interoperability. | | |
| **Existing Recommendations** | | |
| • DOOR[85]<br>• SSSOM[86] | | |
| **Stakeholder** | | |
| Practitioner, Community | | |

| Optional | P-Rec. 13: Crosswalks, mappings and bridging between semantic artefacts should be documented, published and curated | R1.2, R1.3 |
|---|---|---|

[84] OntoUML https://ontouml.readthedocs.io/en/latest/intro/ontouml.html
[85] DOOR http://oro.open.ac.uk/24326/1/keod9.pdf
[86] SSSOM: https://github.com/mapping-commons/sssom



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Description |
|---|
| Mappings, crosswalks, content negotiations, and semantic bridges[87] discussed in the previous recommendation (P-Rec. 12) should be made publicly available to allow the reuse by others. Mappings are semantic artefacts and therefore should be shared and published in semantic repositories, following the recommendations for such artefacts (for example with respect to GUPRI, metadata, licence,...).<br><br>Sharing these resources in a standardised way will improve interoperability. The main requirement is for such mappings to be machine readable for reuse purposes, and to be described with minimum metadata for human interaction. |
| **Existing Recommendations** |
| SSSOM[88]: |
| **Stakeholder** |
| Practitioner, Repository, Community |

## 2.5 FAIR Principles Coverage

Each of the recommendations proposed above are directly aligned with one or more individual FAIR principles. The table 2 below shows this alignment.

| | FAIR Principles | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P-Rec | F1 | F2 | F3 | F4 | A1 | A1.1 | A1.2 | A2 | I1 | I2 | I3 | R1 | R1.1 | R1.2 | R1.3 |
| 1 | ■ | | | | | | | | | | | | | | |
| 2 | ■ | | ■ | | | | | | | | | | | | |
| 3 | | ■ | | | | | | | | | | ■ | ■ | ■ | ■ |
| 4 | | | | ■ | | | | | | | | | | | |
| 5 | | | | ■ | ■ | ■ | | | | | | | | | |
| 6 | | | | ■ | | | | | | | | | | | |
| 7 | | | | | | | ■ | | | | | | | | |
| 8 | | | | | | | | ■ | | | | | | | |

[87] These are all mediators in the language of the RDA Brokering Interest Group

[88] https://github.com/mapping-commons/sssom



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | | | | | | | | | | | | | | | | |
| 10 | | | | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | | | | |
| 12 | | | | | | | | | | | | | | | | |
| 13 | | | | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | | | | |
| 16 | | | | | | | | | | | | | | | | |
| 17 | | | | | | | | | | | | | | | | |

*Table 2: alignment of the recommendations with the individual FAIR principles*

# 3. Recommendations Beyond the FAIR Principles: Best practices for Semantic Artefacts

In this section, recommendations are listed that do not apply to a particular FAIR principle but contribute to the improvement of Findability, Accessibility, Interoperability and Reusability of the semantic artefacts

## 3.1 List of Best practices

| BP1: Use a common naming convention for concept/class and relations |
|---|
| **Description** |
| Concept/class and relations composing semantic artefact are associated with a human readable label which is a character string. This string can become complex and include several associated words (e.g. Hyperplastic and giant kidney**).** There are multiple conventions for naming semantic artefact elements such as CamelCase or the conventions proposed by OBO Foundry and Industry Ontology Foundry. Unfortunately these existing conventions/recommendations are not harmonised which leads to the need for a search engine or an automated mapping service to comply with the different conventions/ recommendations. This hampers both searching capabilities and automated mappings.<br><br>This recommendation for Best Practice emphasises the need to define a common unique naming convention by the community of practitioners. |
| **Existing Recommendations** |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| ● OBO Foundry - Principle 12 (Schober et al., 2009)[89]<br>● IOF - Principle 11[90] |
| **Stakeholders** |
| Practitioners, Community |

| BP 2: Use an Ontology Naming Convention |
|---|
| **Description**<br><br>Semantic artefacts often have a human readable name associated with an acronym. The name provides information about the general topic covered by the semantic artefact while the acronym can be used as prefix to create URI-based GUPRI for concepts and terms (see OBO Foundry naming convention). The governance of these names is managed by organisations such as the OBO Foundry or the Industry Ontology Foundry. As of now, there is no widespread consensus on the naming of semantic artefacts and the use of acronyms/ prefixes, which leads to ambiguities and non-uniqueness (see Goldfarb and Le Franc, 2017). The community of practice should consider addressing this issue by defining a common governance model for semantic artefacts and possibly a unique organisation that would guarantee the uniqueness of acronyms and names. |
| **Existing recommendations**:<br><br>● OBO Foundry - Principle 3[91]<br>● Industry Ontology Foundry[92] |
| **Stakeholder**: Community |

| BP 3: Use defined ontology design patterns |
|---|
| **Description**<br><br>To foster interoperability, semantic artefacts should be designed with defined design patterns whenever relevant and possible. These patterns should be documented and published as a resource for practitioners following the example of OntologyDesignPatterns.org[93] that focuses on OWL design patterns. |

[89] OBO Foundry principle 12 http://www.obofoundry.org/principles/fp-012-naming-conventions.html
[90] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[91] OBO Foundry principle 3 http://www.obofoundry.org/principles/fp-003-uris.html
[92] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[93] OntologyDesignPatterns.org http://ontologydesignpatterns.org/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| Other domain specific examples include:<br>● DOSDP patterns, for example, to reconcile logical definitions across phenotype ontologies (see here).<br>● OTTR library of reasonable templates[94] |
| **Existing recommendations**: N/A |
| **Stakeholder**: Practitioner |

| BP 4: Create mappings validated by domain experts |
|---|
| **Description**<br>Semantic artefacts and in particular concept definitions vary within and between communities. This diversity of definitions generates semantic ambiguities which hampers interoperability between ontologies. To support such interoperability, explicit mappings should be generated by knowledge experts and validated by domain experts. As stated in P-Rec. 12 & P-Rec. 13, these mappings should be serialised with standard formats and be published. |
| **Existing recommendations**: N/A |
| **Stakeholder**: Practitioner |

| BP 5: Define workflows between different formats |
|---|
| **Description**<br>Semantic artefact can be serialised in various formats (SKOS, RDF, OWL, XML, ...). This diversity of formats makes it complicated to integrate and work with heterogeneous semantic artefacts. Practitioners should describe the particular workflow they used to convert the semantic artefact from one format to another. These workflows could be defined using machine readable mappings and should provide means to get notification on potential 'loss of information' in the conversion process. |
| **Existing recommendations**: N/A |
| **Stakeholder:** Practitioner, Community |

[94] http://tpl.ottr.xyz/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## BP 6: Harmonise and document the methodologies used to develop semantic artefacts

**Description**

Semantic artefacts can be built with very different methodologies depending on the available resources and the expertise of the practitioner. These methodologies should be documented to allow external experts willing to reuse the artefact to assess the relevance and quality of the semantic artefact.

To support interoperability and prevent ill-formed ontologies, the community of practice should work toward harmonising and documenting the various methodologies which could be used as a training resource for newcomers and guidelines for expert practitioners.

**Existing recommendations**:

- OBO Foundry - Principle 8[95]
- IOF - Principle 9[96]

**Stakeholder**: Community

## BP. 7: Interact with the designated community and manage user-centric development

**Description**

A semantic artefact needs to reflect the actual semantic model/conceptualisation embraced and endorsed by the community that uses it. As science and language changes, changes and updates are inevitable in the long run or content drift will happen in the semantics. This has to be acknowledged and managed. As others (re)use semantic artefacts to enable interoperability there has to be even more clear processes for managing the necessary communication and negotiations.

Semantic artefacts have to be regarded as services in their own right, not only service components, especially if they are reused (used by several systems or solutions). Thus ownership as well as continuous user centred development should be ensured.

Interaction and communication with the designated community has to be organised and managed as a process. The community should receive training and guidance in using and developing the semantic artefact.

One example of organising the development and maintenance of vocabularies comes from the TDWG - Biodiversity Information Standards community. The TDWG Vocabulary Maintenance Specification details the categories of changes that can be made to a TDWG vocabulary standard, the mechanisms used to

[95] OBO Foundry principle 8 http://www.obofoundry.org/principles/fp-008-documented.html

[96] IOF Technical Principles https://www.industrialontologies.org/?page_id=87



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| achieve those changes, and the entities that are responsible for shepherding those changes through the process. https://github.com/tdwg/vocab/blob/master/vms/maintenance-specification.md |
| **Existing recommendations**:<br>● OBO Foundry - Principle 9 Users[97]<br>● OBO Foundry - Principle 11 Authority[98]<br>● OBO Foundry - Principle 16 Maintenance[99]<br>● IOF Principle 16[100]<br>● IOF Principle 15<br>● IOF Principle 14 |
| **Stakeholder**: Practitioner |

| BP. 8: Provide a structured definition for each concept |
|---|
| **Description**<br><br>Semantic artefacts are used to annotate data. Concepts are the key elements of the semantic artefacts used by the annotators. Annotators need both a human readable and logical definition to make a decision on using a specific term. Human readable definitions are therefore crucial for the reuse of semantic artefacts. When building a semantic artefact, one the main challenge is to write a structured definition for concepts. There are already quite a few recommendations providing guidelines for writing human readable definitions as OBO Foundry principle 6 and the IOF principle 10. A set of guidelines have been proposed following up a dedicated series of workshops and a survey on the usage of definition (Seppälä et al., 2017). A recent blog post from C. Mungall describe in more detail how to write simple and concise definitions[101]. |
| **Existing recommendations**:<br>● OBO Foundry - Principle 6[102]<br>● IOF - Principle 10[103]<br>● Guidelines for writing definition in ontologies[104] |

[97] OBO Foundry principle 9 http://www.obofoundry.org/principles/fp-009-users.html
[98] OBO Foundry principle 11 http://www.obofoundry.org/principles/fp-011-locus-of-authority.html
[99] OBO Foundry principle 16 http://www.obofoundry.org/principles/fp-016-maintenance.html
[100] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[101] OntoTip: Write simple, concise, clear, operational textual definitions https://douroucouli.wordpress.com/2019/07/08/ontotip-write-simple-concise-clear-operational-textual-definitions/
[102] OBO Foundry principle 6 http://www.obofoundry.org/principles/fp-006-textual-definitions.html
[103] IOF Technical Principles https://www.industrialontologies.org/?page_id=87
[104] Guidelines for writing definitions in ontologies https://philpapers.org/archive/SEPGFW.pdf



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| |
|---|
| |
| **Stakeholder**: Practitioner |

| BP. 9: The underlying logic of semantic artefacts should be grounded on the domain it intends to describe |
|---|
| **Description**<br><br>Semantic artefacts are developed within research communities and represent a specific and restricted domain of discourse. Reuse of such ontologies by stakeholders outside of the community raises questions regarding the relations between the ontologies and scientific domains. When reusing ontologies, practitioners should strive to choose semantic artefacts with highest precision (e.g. existence of information cardinality and value), the most well defined documentation including information regarding cardinality and value type when applicable.<br>Whenever the semantic artefact is reused, it should be extended in granularity depending on use-case. This implies that the reuser should adhere to the same design principle. |
| **Existing recommendations**: N/A |
| **Stakeholder**: Practitioner |

| BP. 10: Define a set of governance policies for the semantic artefacts |
|---|
| **Description**<br><br>Semantic artefacts are living entities that are undergoing changes through their lifecycle. As an example, when used these artefacts are becoming part of a data service and therefore changes and updates should be published to warn users that the data service should be updated. It is therefore crucial to have well identified governance policies for the semantic artefact. These policies should be available in human readable format but also whenever possible in machine actionable format to allow the automation of change propagation. They should cover the various aspects of the semantic artefact life cycle i.e. versioning policy, deprecation policy, contribution policy. An example of such policies would be the TDWG Vocabulary Management Standard[105]. |
| **Existing recommendations**: N/A |
| **Stakeholder**: Practitioner and Repository |

[105] https://www.tdwg.org/standards/vms/



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

## 3.2 Analysing the Best practices landscape

There are several best practice recommendations and articles about good practices for creating, developing, maintaining, sharing and re-using semantic artefacts. Existing registries for semantic artefacts have needed internal guidelines for governing the content and making it usable for the original purpose of the registry, such as the OBO Foundry Principles (Smith et al., 2007) and Industry Ontology Foundry Technical Principles (Kulvatunyou et al., 2018). However, governance models and guidelines for FAIR semantics are needed for not only creating new FAIR semantic resources, or improving maturity of existing semantic artefacts in a repository, but also for publishing legacy resources in a FAIR format making them discoverable and re-usable across disciplines or themes.

This need for aligning the principles has been addressed in several recent articles (Garijo et al., 2020; Jackson R. et al., 2021; Karray et al., 2021; Cox et al., 2021; Amdouni et al., 2021).

Our best practice recommendations suggest that the community of practice should consider defining a common governance model for semantic artefacts, including their URI design, naming conventions, and versioning at a general level. These issues have been also recognised quite recently by the practising communities, and more detailed BP guidelines, suggestions, examples and tools have been presented in recent literature.

So far, most detailed and perhaps mature best practice guidelines or principles have been developed for existing ontology or vocabulary registries or repositories that serve a particular purpose (theme, domain or scientific discipline). These registries are actively governed and to be accepted and published in such a registry, semantic artefacts go through a review process that is normally carried out by the community representing the same purpose. However, publishing and governance of independently developed (single purpose) semantic artefacts in these registries is laborious and this has hindered opening them for wider user communities. Present research is clear in concluding that existing ontologies still suffer from a lack of interoperability. (Karray, M., et al. 2021)

Cox et al (2021) suggest a stepwise approach e.g. ten simple rules that support converting a legacy vocabulary into a FAIR vocabulary, which can be used for unambiguous data annotation. In turn, this increases data interoperability and enables data integration, which is essential for addressing global challenges such as environmental sustainability, and pandemic and natural disaster response. Various pathways may be followed to publish the FAIR vocabulary, but particular emphasis is given to the goal of providing a globally unique resolvable identifier for each term or concept. They also suggest further steps towards broader interoperability that may be considered in the future should include: 1) relationships to terms and definitions in other FAIR vocabularies 2) patterns for re-use of terms from and subsets of existing FAIR vocabularies, 3) supplementation of generic SKOS/OWL encoding with domain-based elements and axiomatization and 4) rules for maintenance.

Jackson, R. et al (2021) recognise that the need for consistent governance models and principles are compounded when considering the fact that many applications require using 'combinations' of



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

ontologies. If ontologies are constructed using different principles, they will not work together in a modular, interoperable, and coherent way. So far, the OBO foundry's review process has been relying on volunteer effort and manual work. In the article they describe efforts to operationalize the OBO Foundry principles by making an automated review process possible. Working closely with stakeholders across OBO, they have refined the principles, codifying them into operational tests that can be executed automatically at regular intervals, thus helping the governance process significantly. They also recognise the need to provide OBO resources in a FAIR format for the outside world, and together with the stakeholders, they agreed on how the licence should be stated and decided on the use of the widely accepted Dublin Core Terms (Weibel S., Kunze J., Lagoze C. et al. 1998) 'licence' property ('dcterms:licence') in the ontology file metadata in addition to a declaration of the licence in the OBO registry entry. These conventions allow checking for the presence of an 'Open' licence computationally, in both the ontology file itself and the information contained in the OBO registry.

As Jackson et al. analyse the OBO registry principles, Garijo & Poveda (2020) operate on a generic level and suggest concrete solutions and tools as they raise the Accessible Ontology URI Design, Generating Reusable Ontology Documentation, and Ontology Publication on the Web to be the main themes of their best practice recommendations. The recommendations provide justification and concrete examples of using opaque URIs for classes and properties, semantic versioning and recommendations for creating re-usable ontology documentation. The recommendations about documentation include considerations of minimum metadata (recommended and optional), ways and tools to visualise the ontology model and also publication frameworks. The recommendations also distinguish two main categories of metadata in an ontology: the metadata associated with the ontology itself and the metadata associated with its elements (classes, object properties, datatype properties and individuals), which is inline with our recommendations. While the FAIRsFAIR best practice recommendations go beyond FAIR, they emphasise the need for alignment / harmonising methods between the domain specific registries. Garijo & Poveda (2020) also operate on a general level, but they address this issue from a different angle. They give concrete, but general recommendations on how to publish FAIR ontologies on the Web while following the Linked Data principles. A distinct feature of their guidelines is that they illustrate how to carry out the recommendations with an example ontology and pointers to usable tools developed by the Semantic Web community in the last decade.

Despite the recent interest in the open science and Semantic Web communities, a quantitative evaluation method to assess and score the level of FAIRness of ontologies or semantic artefacts has been missing until recently. Amdouni & Jonquet (2021) present an integrated quantitative assessment grid for semantic resources and propose candidate metadata properties –taken from the MOD (Metadata for Ontology Description and publication) – to be used to make a semantic resource FAIR. The grid distributes 478 credits to the 15 FAIR principles in a manner which integrates existing generic approaches for digital objects (i.e., FDMM, SHARC) and approaches dedicated to semantic resource or artefact (i.e., 5-stars V, MIRO, FAIRsFAIR, Poveda et al.). The credits of the grid can then



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

be used for implementing FAIRness assessment methods and tools. They conclude that from a semantic Web perspective, the results obtained emphasise the need for the establishment of agreement about a set of core metadata ontology description, a federation model for ontologies regarding repositories and search engines, clear ontology and metadata ontology perseveration strategies within endpoints, mechanisms for references qualification, and best practices to document and communicate ontologies. Their work shows that qualifying the degree of FAIRness of a semantic resource or even comparing it with other semantic resources necessarily implies the use of a metric delimiting the range of values for each qualification (e.g., not FAIR, FAIR, or FAIRer). In that context, the proposed integrated quantitative grid allows defining thresholds and using a metric and thresholds is a first required step in making the FAIRness assessment task machine actionable and enable the development of automatic FAIRness assessment tools.

# 4. From recommendations to practical implementation

One of the most important inputs we received during the various workshops and consultations was the need to enrich the recommendations with concrete and practical solutions for implementing them. The purpose of the recommendations was to just provide general guidelines, leaving freedom of implementation. However, as some recommendations require community work, we chose for this third version of the document to include more practical aspects. First, we are proposing an initial generic service architecture for leveraging FAIR Semantic artefacts that we extended with recommendations for the design of the EOSC infrastructure. Second, we are describing our effort to propose a community driven minimum metadata schema for FAIR Semantic Artefacts.

## 4.1 Designing an optimal service architecture for leveraging FAIR Semantic artefacts

The implementation of FAIR principles opens the door to automation and better integration of the tools for working with Semantic artefact. The goal is to facilitate the work with semantic artefacts. Therefore, we chose to focus on one particular stakeholder: the semantic artefact developer/manager who will create, use and reuse semantic artefacts.

We developed a use-case centred around these roles. In the end, this person will create semantic artefacts together with a community of users within a scientific domain. These semantic artefacts should reuse other existing and related artefacts in order to reduce semantic ambiguity for machines. In order to support the reuse of their semantic artefacts efficiently, semantic artefact developer/manager will need to publish them into a repository which will provide access to the semantic artefact content through various clients. When reusing semantic artefacts, semantic artefact developers/managers first need to find them, assess their quality and investigate existing mappings and/or create new ones. These mappings can, then, be reviewed by the community who evaluate their potential for reuse beyond the use cases for which they have been defined.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Based on this description, we identified, for each of these actions, 8 necessary services and some of their requirements. We focused in particular on one of the key pieces of this FAIR Semantics infrastructure: the semantic repository. The outcome of this theoretical work is described below. The user stories can be found in the Annex. Please note that this work is exploratory and is proposed as a first suggestion to be further discussed and refined by various working groups related to semantic interoperability such as the EOSC Semantic Interoperability Task Force (see Future Work).

We established a simple diagram to articulate these services shown below in figure 4.

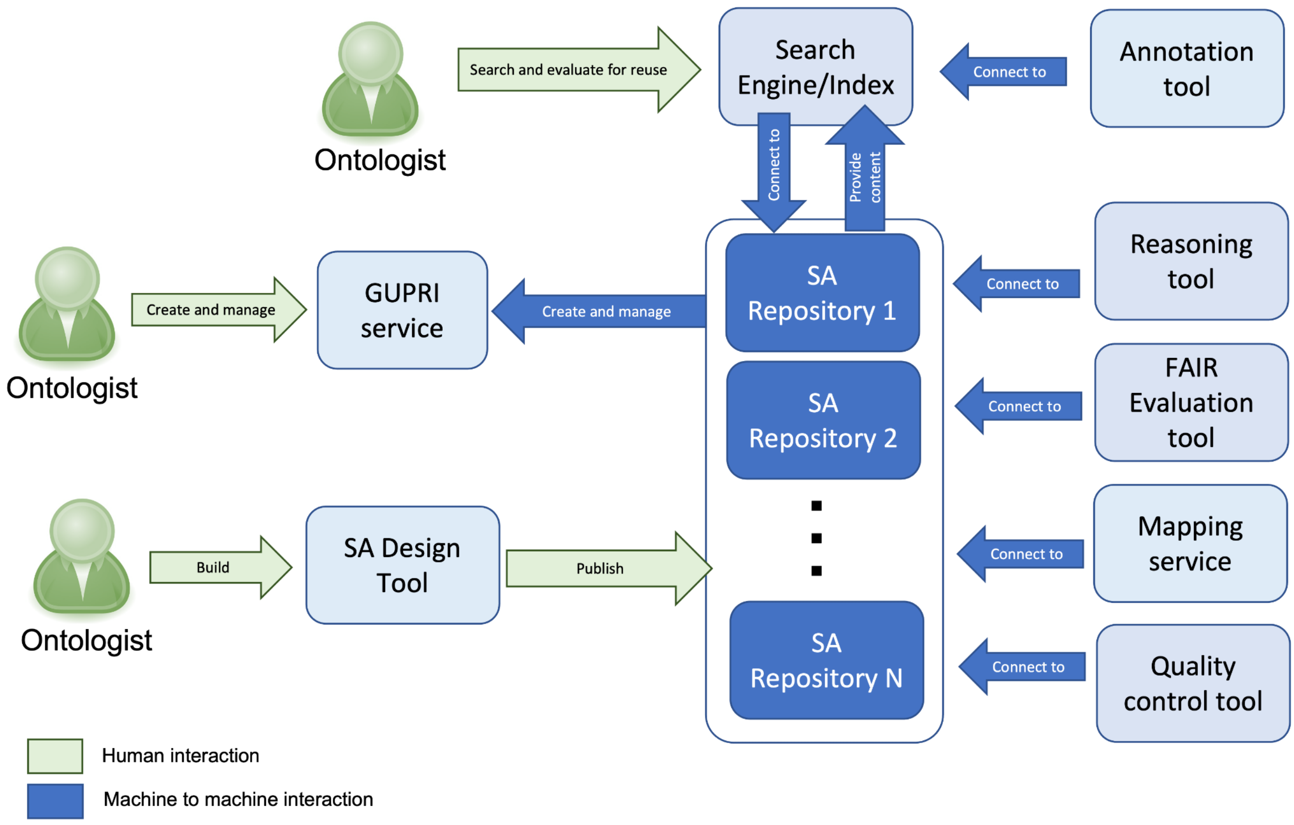


Figure 4: FAIR Semantic architecture diagram

The services presented here are further described below.

- **A GUPRI service.** This service should provide means to create and manage Globally Unique Persistent and Resolvable Identifiers as defined in the EOSC PID policy (Hellström et al., 2020). As semantic artefacts are often Web-based, we are recommending to use Persistent URLs for identifying the semantic artefact and its content. By nature URLs/URIs are Globally Unique and Resolvable but not necessarily persistent, as we can see with the many dead links in the Web. The service should therefore guarantee ideally a long-term maintenance of these identifiers. As identifier policies are tightly dependent on organisation policies and also on the usage of the identifiers, the service should define persistence policies for the



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

provided identifiers. These policies should be both published in human and machine readable formats. Making the policies machine readable will allow access to this information within tools when selecting ontologies for re-use. This Web-based GUPRI service should be considered as one of the key services of EOSC infrastructure. Governance policies could thus be defined at a global level rather than at community level which will reduce complexities.

- **A service/tool to create semantic artefacts.** Various tools already exist for building semantic artefacts, such as the popular Protégé[106] (Musen et al., 2015). These tools are the main interface for the semantic artefact developer/manager and therefore they should be better integrated with the proposed architecture to provide additional features to their users. In particular, it would be useful to enforce metadata capture by providing agreed upon metadata schema such as the minimum metadata schema we are presenting in the next section.
- **A service to store and publish semantic artefacts.** Several technologies exist to store,publish and search for semantic artefacts (e.g. OntoPortal virtual appliances[107], SKOSMOS[108], EBI-OLS [109], etc.) as well as standalone services. This service has been identified as a key component of the EOSC Semantic Interoperability Framework (Corcho et al., 2021). We see many of these services set up with domain specific focus. These repositories should offer a simple interface to upload and publish, visualise and interact with semantic artefacts and their content through a REST API. As we understand that redesigning APIs has a cost for these resources, we would like to recommend using existing frameworks for federating APIs such as SmartAPI approach[110]. This approach relies on the OpenAPI framework for documenting API for humans and extends it to become a machine readable documentation of the API. By harmonising the API landscape with machine readable API descriptions it becomes possible to easily interconnect the semantic repositories together in order to build a search engine as described below. Another requirement for the repository should be to provide a machine readable version of their metadata schema which will ease the development of generic clients able to manipulate their content. By doing so, they will comply with the FAIR Data Point specification[111]. Finally, semantic repositories should be integrated with the GUPRI service to provide a GUPRI when necessary.
- **A search engine.** To foster reuse of existing semantic artefacts either as whole or as part, semantic artefact developer/manager should be able to access the wealth of semantic artefacts through a dedicated search engine which would aggregate the various existing semantic repositories. Although the existing semantic repositories offer such search engines, they are solely restricted to their content and require semantic artefact developer/manager

[106] https://protege.stanford.edu/
[107] https://ontoportal.org/the-ontoportal-virtual-appliance/
[108] https://skosmos.org/
[109] https://www.ebi.ac.uk/ols/index
[110] https://smart-api.info/
[111] https://specs.fairdatapoint.org/



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

to search through multiple repositories to collect the relevant information. The main barrier to create such search engines is the diversity of repository APIs and metadata schemas. By implementing the requirements provided earlier, it would be possible and manageable to create a semantic artefact search engine that would collate as many sources as possible. semantic artefact developer/manager could then search in one place existing concepts/terms/semantic artefact/relations that could be either of use or that would overlap with concepts from other ontologies, similarly to the discontinued Swoogle[112] (Ding et al., 2004). This search engine should leverage an index created from the harvesting of the repositories. This index could then be integrated into existing tools to provide fast access to concepts and relations. The index should not contain the complete copy of the repositories but provide links to collect extra information.

- **A service/tool to assess the quality and FAIRness of semantic artefact.** In order to decide whether to reuse or not a semantic artefact, the semantic artefact developer/manager should be able to evaluate the quality of the semantic artefact. For this purpose, semantic repositories should be connected or integrated with existing tools for quality assessment such as OOps![113] (Poveda-Villalón et al., 2014) or FAIRness assessment such as O'FAIRe (Amdouni et al., 2021) and FOOPS! (Garijo et al., 2021). This is actually the case for the O'FAIRe assessment tools which is integrated with AgroPortal and leverages the usage of MOD to describe the ontologies in AgroPortal. Existing FAIRness evaluation tools more oriented towards data such as F-UJI[114] or the FAIRSharing evaluator could be extended to support the evaluation of semantic artefacts.
- **A service to create, access and reuse curated semantic mappings.** When aligning with other ontologies, the semantic artefact developer/manager should be able to create and reuse existing mappings between semantic artefacts. These mappings should be reviewed by the community and should therefore provide the necessary functionalities. SEMAF proposes a framework for developing such services (Broeder et al., 2021).
- **A service for annotating data with semantic artefacts.** Once the semantic artefacts have been built, the semantic artefact developer/manager should be able to benchmark them with real data and use them for the specific use-case. For this, the semantic artefact developer/manager should be provided with tools for annotating data. Tools for data annotation already exist such as B2NOTE (Kulhanek et al., 2019) and its extension Semaphora (Molloy et al., 2021). This EOSC service should leverage the semantic index supporting the search engine to provide support for semantic tagging.
- **A service for reasoning over the semantic artefacts.** An additional service to be considered should be a reasoning service which would allow evaluating semantic and logical inconsistencies but also inferring new knowledge from multiple semantic artefacts when

[112] https://ebiquity.umbc.edu/project/html/id/53/Swoogle
[113] http://oops.linkeddata.es/
[114] https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

these artefacts have a logical framework. This service should be able to access the semantic artefacts directly from the semantic repositories. This service is nice-to-have, but would contribute to maintaining semantic artefacts quality and reusability, and it would be a basis for applying AI in a semantic space.

This presented simple architecture extends the proposed infrastructure for the EOSC semantic interoperability framework. By promoting further integration of these services, we will be able to offer an improved working environment for semantic artefact developers/managers.

Another major practical aspect of the implementation of the FAIR Semantics recommendation is the use of common minimum metadata schema for describing and publishing semantic artefact. In the next section, we present our work and the resulting minimum metadata schema.

## 4.2. Defining a minimum metadata schema for FAIR Semantic Artefacts

In order to properly follow the FAIR principles, semantic artefacts need structured and interoperable metadata. Interoperability happens at different levels and at different degrees. For instance, by agreeing on a common metadata presentation format (e.g., RDF, JSON, XML) we improve interoperability by allowing the development of client applications that would need to read/parse only one presentation format. On top of this, by agreeing on a minimal metadata schema for semantic artefacts, we would guarantee that every time an application encounters the metadata of a semantic artefact, it would know that minimally certain information would be available. The recommendation P-Rec. 3 states that semantic artefacts and their content *must* be described using a common minimum metadata schema.

During the course of the FAIRsFAIR project work has been done on defining a common metadata description for SAs. This work was initiated with a few members of the RDA VSSIG TAG on metadata led by Clement Jonquet. It has been taken further towards a DCAT extension for representing SAs and a short list of metadata fields to describe them. The goal of the common work has been to propose a minimum metadata description for FAIR SAs and express it in an machine actionable, interoperable way, thus setting up a base threshold on FAIRness. This approach does not restrict the usage of more metadata elements for richer metadata description as required by the FAIR Principles.

Data Catalogue Vocabulary is an RDF vocabulary designed to facilitate interoperability between open data catalogues published on the Web. By using DCAT to describe datasets in catalogues, publishers increase discoverability and enable applications to consume metadata from multiple catalogues. The DCAT Application Profile is used by Data Portals in Europe (DCAT-AP) and it is also recommended in the EOSC Interoperability Framework (Corcho et al., 2021). The DCAT Model is shown in the figure 5 below. DCAT offers a basis for a flexible way to describe semantic artefacts as digital objects, but there is a need to define a specific extension for semantic artefacts.



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

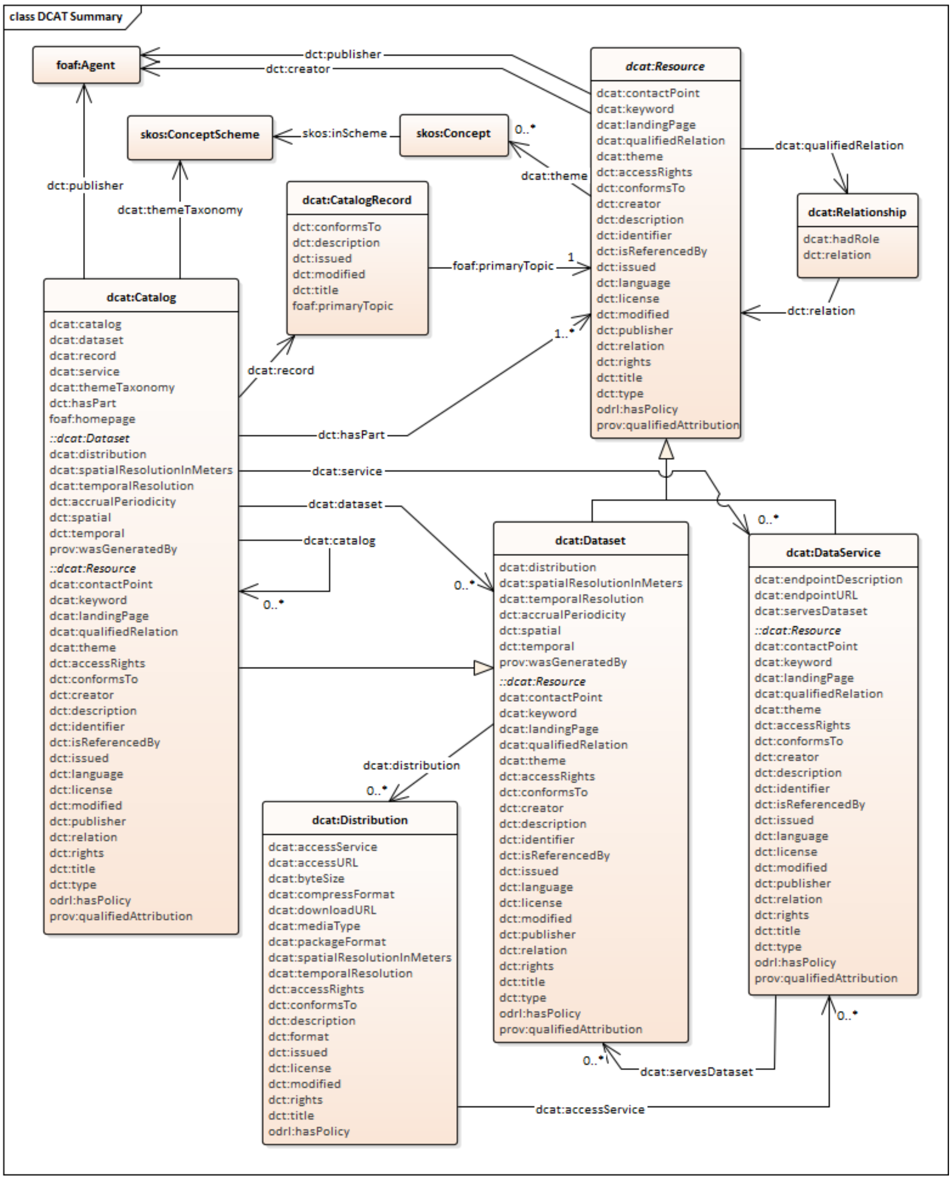


Figure 5: Overview of the DCAT model from [https://www.w3.org/TR/vocab-dcat-2/](https://www.w3.org/TR/vocab-dcat-2/)

DCAT provides a high level description for representing collections of semantic artefacts but does not integrate specific metadata fields to describe the semantic artefact themselves. For this purpose, several models have been proposed such as OMV, VoID or DOOR to describe either the ontology itself or its relations. The Metadata vocabulary for Ontology Description and publication also called MOD has been proposed by Dutta et al. (Dutta et al., 2017). MOD 1.4 has been designed by reviewing in total 23 standard existing metadata vocabularies (e.g., Dublin Core, OMV, DCAT, VoID) and selecting relevant properties for describing ontologies. MOD 1.4[115] proposes in total 146 properties to serve both as (i) a vocabulary to be used by ontology developers to annotate and

[115] https://www.isibang.ac.in/ns/mod/index.html



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

describe their ontologies, or (ii) an explicit OWL ontology to be used by ontology repositories/ libraries to develop and offer semantic descriptions of ontologies as linked data. During the course of the project, we have been working extensively with MOD authors to create a new version of MOD which includes the alignment with DCAT.

As most of the SA have scarce or no descriptive metadata at all, imposing the use of the full MOD model would not be possible (146 relevant metadata fields). A large part of the work initiated within the RDA VSSIG was to identify a subset of metadata fields from MOD. This subset should be then used to extend the DCAT model with metadata elements specific to Semantic Artefacts. Once this preliminary work was done, we worked toward collecting feedback from the community of experts. For this purpose, we organised an expert workshop, held on Zoom on 4 June 2021 at 9:30-13 CEST.

The workshop was organised with two specific aims: 1) presenting and discussing the outcome of the RDA VSSIG task group i.e. the alignment with DCAT and the use of a minimum set of metadata fields from MOD; 2) collect feedback on the identified metadata fields i.e. should they be mandatory, recommended or optional in order to define a DCAT Application Profile (DCAT-AP) for Semantic Artefact. This profile should contain a subset of the existing DCAT metadata field as well as an extension which includes MOD metadata fields.

To guide the discussion, we proposed a simple use-case (searching for an ontology), presented prior to the interactive second part of the workshop. It's purpose was to describe the usage of DCAT as a common metadata model and to define specific metadata elements describing semantic artefacts so that it would enable creating a search engine which would allow retrieving semantic artefacts hosted within various heterogeneous and distributed repositories.

To specify the requirements the participants produced data on whether relevant properties from the classes dcat:Resource, dcat:Dataset, dct:Catalogue, dcat:Service, dcat:Distribution and mod:SematicArtefact and mod:SemanticArtefactDistribution should be optional, recommended or mandatory and also the relevance for enabling the FAIR principles to be promoted. To collect these inputs, we used Mentimeter. The mentimeter survey thus produced data on how experts assessed the relevance of the different elements of MOD2 as an extension of DCAT2.

At the beginning of the workshop all attendees were made aware that to participate in the interactive session, in the second half of the event, an understanding of data modelling and a good familiarity with DCAT was needed. During this part of the workshop participants were asked only to share responses for which they considered themselves to have sufficient expertise and awareness to make an informed contribution. They were guided through this process by the FAIR Semantics team who answered specific questions relating to DCAT, MOD and FAIR as and when necessary. The



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

participants in the workshop were engaged to provide their expertise using the Mentimeter tool, and the facilitator of the session provided troubleshooting for engaging with the platform.

For each of the relevant DCAT and MOD classes a list of properties were presented using the 'Sliders' function. This allowed participants to plot each property along an axis with three points (1) Mandatory; (2) Recommended; and (3) Optional. Although the 'Multiple Choice' function may have been a better fit for collecting this type of information, due to the limitations of the tool it would have required at least 120 individual questions to collect the required data whereas using the 'Sliders' function meant that properties could be bundled in groups of 10 making the process less time consuming and more user-friendly. A similar workaround was used for performing the mapping of properties against each of the principles of FAIR - Findable; Accessible; Interoperable; and Reusable. Using the 'Grid' functionality participants were again asked to plot each property this time using two axes: Findable <-> Accessible; and Interoperable <-> Reusable. Participants were instructed to use only one of the two sliders available per property to identify the principle most closely related to that property. A visual example and written instructions were shared with participants via presentation and also within the body of the workshop agenda and collaborative notes document (See ANNEX 1).

It should be noted that despite the high level of support and guidance, adapting Mentimeter functions in this way did somewhat impair the ability of all participants to share their expertise. Those participants who had less familiarity with Mentimeter or are less comfortable working in a multi-screen and multi-platform environment still struggled with this task. Questions and comments were made during the session about the fact that the 'F and A', and 'I and R' of FAIR are not counterpoints on a spectrum indicating that they had not fully understood the activity. This provides evidence that even when using high-end online engagement platforms with highly technical people there are limitations to what can be achieved in an online environment, and that a lot of work is still needed in this area both by users, end-users, and developers of these systems.

The results were collected into a spreadsheet for evaluation. The inputs were the number of votes for each endorsement level mandatory-recommended-optional. From them, we calculated the percentage of votes for each option. The option with the highest percentage was then selected. In cases where two options voted the same, the third was taken into account, e.g. :

- 46.43% mandatory, 46.43% recommended, 7.14% optional -> recommended, as 7.14% also voted for optional
- 25.00% mandatory, 37.50% recommended. 37.50% optional -> recommended, as 25% also voted for mandatory

As an auxiliary informative metric, we computed the consensus of the voting using the following formula: consensus = (0.333 + percentage of votes for the winning option − sum of percentages of



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

the non-winning options) / 1.333. The different fields considered in this first version of the minimum metadata schema as well as the voting results are provided in the annex.

Once the results of the votes were analysed, we created machine readable representations of the minimum metadata schema to support its use and testing. We created three representations: RDF, OWL and SHACL that available for testing and comments on github [116]..

The RDF description has been elaborated according to the DCAT Application profile for data portals in Europe (DCAT-AP) version 2.0.1[117]. It describes the classes and attributes in the form of annotations (rdf:Description). The benefit of this format is the structural alignment with DCAT-AP, however it exhibits limited machine-actionability -- there is no structural information and also the endorsement level (mandatory/recommended/optional) is contained just in the text description.

```
<!-- dct:LicenseDocument -->
    <rdf:Description rdf:about="http://purl.org/dc/terms/LicenseDocument">
       <vann:usageNote xml:lang="en">Mandatory class. A legal document giving official permission to do something with a resource. </vann:usageNote>
       <dct:identifier>dct:LicenseDocument</dct:identifier>
       <rdfs:label xml:lang="en">Licence Document</rdfs:label>
       <rdfs:isDefinedBy rdf:resource="http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=terms#LicenseDocument" />
       <rdfs:comment xml:lang="en">A legal document giving official permission to do something with a Resource.</rdfs:comment>
    </rdf:Description>
```

Example of a property description in SemanticDCAT-AP

To enhance the machine-actionability of the representation, we created an OWL representation. The properties are mapped to their respective classes using OWL Object Properties construct (OP). OPs are annotated by a semantic representation of an endorsement level which allows to identify the mandatory, recommended and optional properties. Figure 6 below shows an example of property mapping in Protégé.

[116] https://github.com/FAIRsFAIR/SemanticDCAT-AP
[117] https://github.com/SEMICeu/DCAT-AP



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

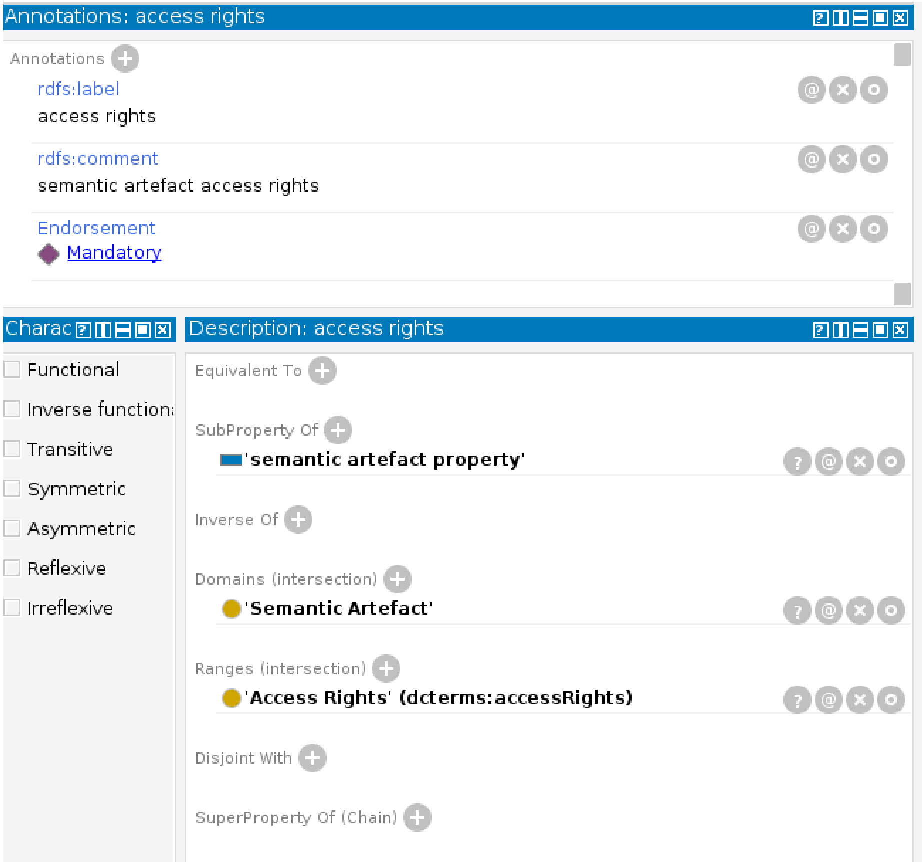


Figure 6: Example of a property mapping in the OWL representation in Protége

This enables effective retrieval of all properties of Semantic Artefact and Semantic Artefact Distribution and their endorsement using SPARQL, as shown in the figure 7 below.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

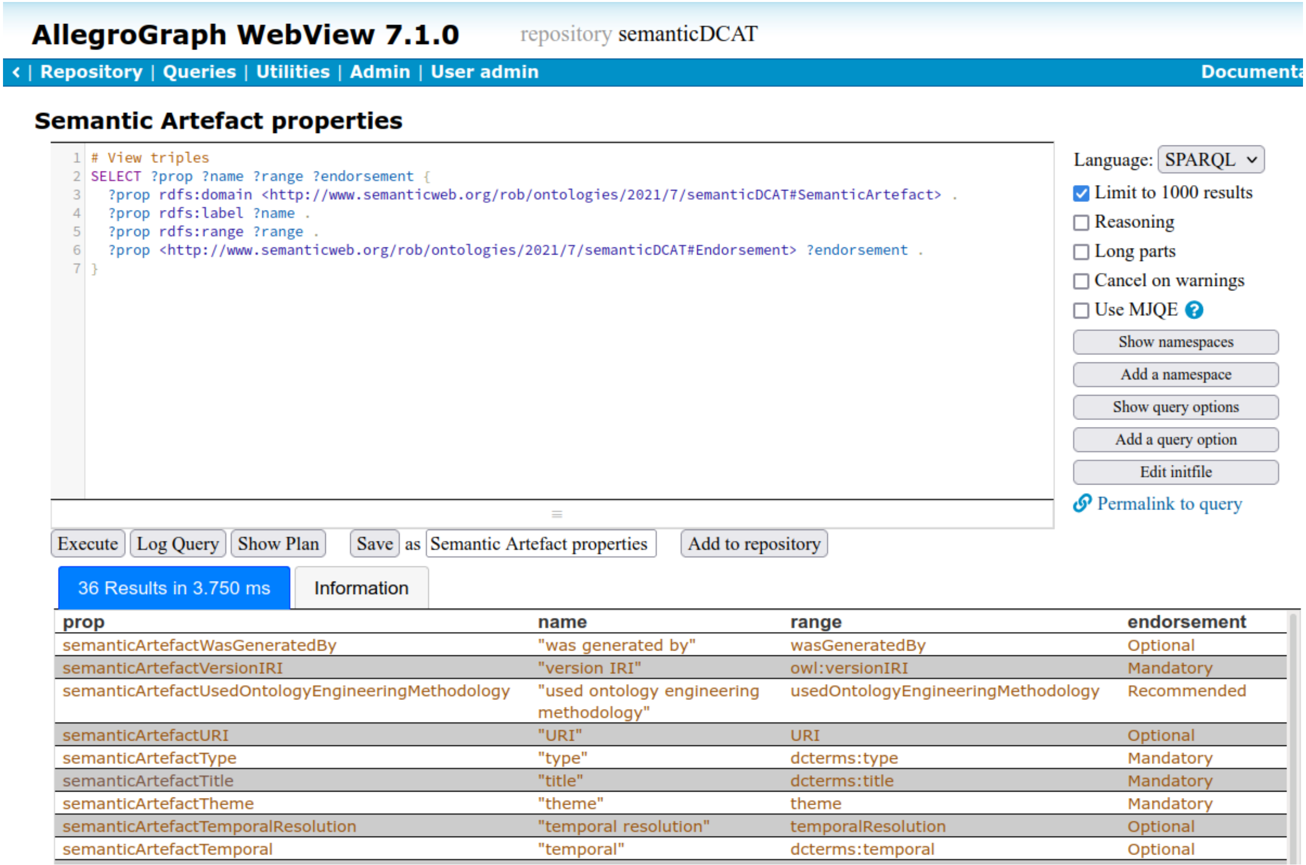


Figure 7: Example of a SPARQL query populating Semantic Artefact properties from a triple store

Details about a certain property can then be retrieved by subsequent queries as shown in figure 8 below.



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

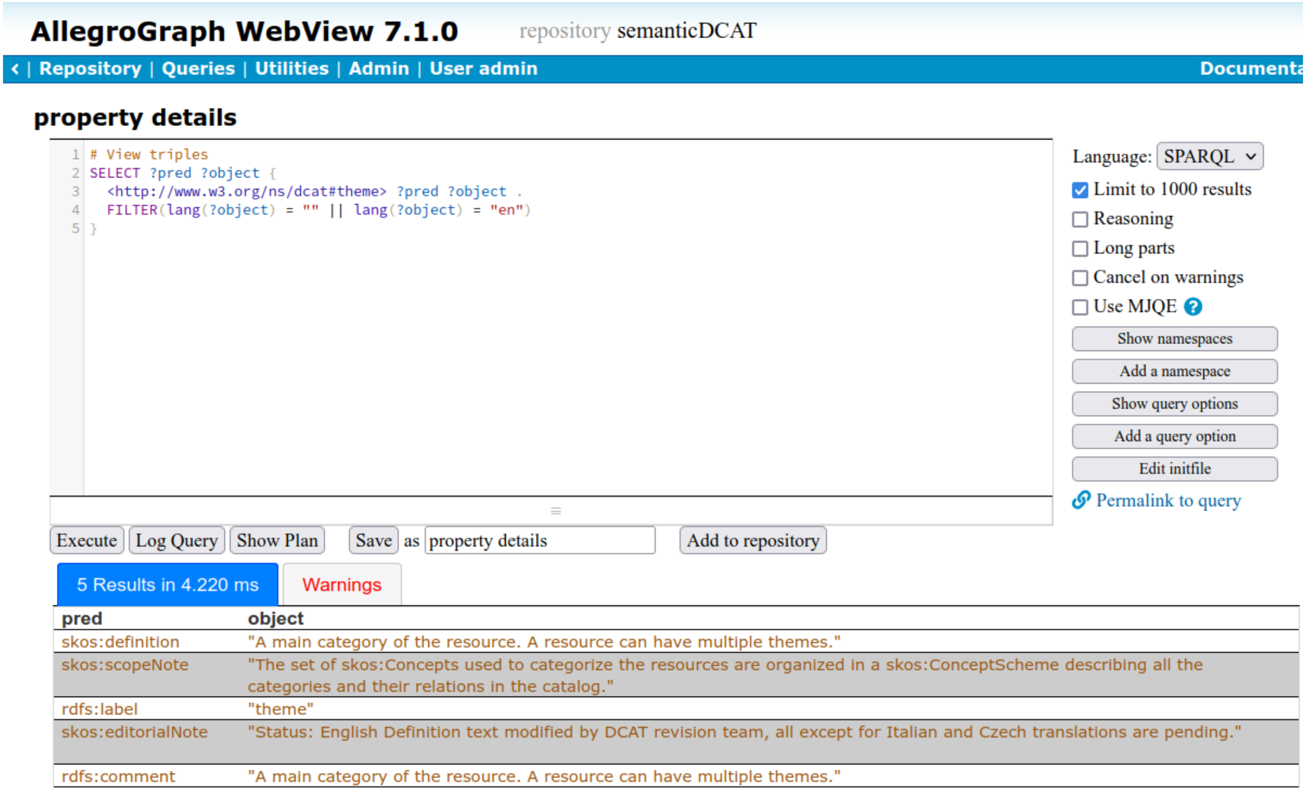


Figure 8: Example of a SPARQL query populating details of a property from a triple store

To provide a way for structural validation of SemanticDCAT ontologies, we also created the Shapes Constraint Language (SHACL) description (included in the repository). This artefact enables to verify compliance of an ontology with SemanticDCAT. Moreover, it can be directly uploaded into a FAIR Data Point (FDP) Reference Implementation to facilitate metadata input thanks to user input annotations.

Based on this work, we developed a Proof-of-Concept search engine which leverages the federated FAIR Data Space developed by the EOSC Pillar project. This technological stack uses SmartAPI for documenting the repository APIs, offers a mapping interface to align the repository metadata schema to the minimum metadata described below. The repository metadata is then published in an instance of the FAIR Data Point reference implementation. This work has been done in the context of the T2.3. in collaboration with EOSC Pillar. For more details, please read the deliverable D2.9.

We used the F2DS to create SmartAPI descriptions representing the API calls necessary to collect all the metadata of all the semantic artefact hosted in various semantic repositories. For this Proof of Concept, we focused on three repositories: Bioportal, AgroPortal and the NERC Vocabulary Server.

The machine readable representations were then used to create mapping templates to convert the repository metadata schema into the DCAT-AP for Semantic Artefact. To create these mapping templates, we tuned the mapping User Interface to upload and visualise the DCAT profile for



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Semantic Artefact. Once this mapping template was created, the metadata was then converted and published into an instance of the FAIR Data Point reference implementation. The metadata description can be then validated within the FDP using the SHACL description. This work is detailed in the deliverable D2.9 and is still ongoing. The resulting API description, mapping template and the link to the FAIR Data Point instance will be made available on the github repository along with the machine readable descriptions of the DCAT profile for Semantic Artefacts.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

# 5. Final Remarks

This document marks the final iteration of the FAIR Semantic team to provide meaningful recommendations for knowledge experts and semantic tools builders and maintainers to work with FAIR Semantic artefacts. In this version, the recommendations and best practices are stable compared to the previous version showing that an initial consensus has been reached with the community of experts. This consensus needs now to be confronted with the reality and with more practical aspects linked to their implementation. This consensus is shown also by the impact of the recommendations on the creation of practical proposition to implement them (Garijo et al., 2020; Cox et al., 2021), on the evaluation of Semantic Artefact FAIRness (Amdouni et al., 2021; Garijo et al., 2021; Devajaru et al., 2021), on the ongoing discussion within various communities (Biodiversity: Sterner et al., 2021; Disaster: Mazimwe et al., 2021; Marine Biology: Kokkinaki A. et al., 2021; AI: Basereh et al., 2021). In addition, the recommendations are currently evaluated for industrial semantic artefacts in the OntoCommons project and are also considered as relevant material in the context of the EOSC Semantic Interoperability Task Force.

To support the development of more practical recommendations, we worked with the community to define one of the missing pieces for the implementation of the FAIR principles for Semantic Artefacts, a minimum metadata schema. This minimum metadata schema aligned with DCAT to support better interoperability can be used as a threshold for defining minimally FAIR semantic artefacts. This schema should now be further discussed and evaluated by the community of experts.

In parallel, we started to define an initial service architecture that would ease the work with semantic artefact with dedicated requirements for the services so they can leverage and comply with the FAIR recommendations.

This work has been the starting point for discussing the harmonisation of the practices in the semantic web community. The emergence of new Best practices is a great example of the impact of this work. However, the existence of multiple Best practices adds more confusion for the community of practice. One of the key challenges that should be addressed in the future is the alignment of these best practices toward one unique set of practices which will ease the implementation of the FAIR principles and alleviate potential interoperability issues arising from the different perspectives.

The work initiated in this project should be continued in another setting and should involve as many communities as possible to reach convergence and to provide an extensive list of concrete examples of implementation. As a large part of the work presented in this document has been done in collaboration with the RDA Vocabulary and Semantic Service Interest Group, the work should be pursued within a dedicated Working Group. A recent survey of the VSSIG members revealed a strong interest in contributing to this Working Group. This working group should be built together with the other initiatives involved in establishing common best practices for FAIR Semantic artefacts.



FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

# 6. References

*Note: Sources like services, web resources, policy documents and specifications are given in footnotes as full http links to enable easy access while reading this document.*


Amdouni E. & Jonquet C. (2021) FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies DOI: 10.13140/RG.2.2.23265.84321

Basereh, A. Caputo and R. Brennan, "FAIR Ontologies for Transparent and Accountable AI: A Hospital Adverse Incidents Vocabulary Case Study," 2021 Third International Conference on Transdisciplinary AI (TransAI), 2021, pp. 92-97, doi: 10.1109/TransAI51903.2021.00024.

Broeder, Daan, Budroni, Paolo, Degl'Innocenti, Emiliano, Le Franc, Yann, Hugo, Wim, Jeffery, Keith, Weiland, Claus, Wittenburg, Peter, & Zwolf, Carlo Maria. (2021). SEMAF: A Proposal for a Flexible Semantic Mapping Framework (1.0). Zenodo. https://doi.org/10.5281/zenodo.4651421

Corcho, Oscar ; Eriksson, Magnus ; Kurowski, Krzysztof ; Ojsteršek, Milan ; Choirat, Christine ; Sanden, Mark van de ; Coppens, Frederik, 2021, “EOSC Interoperability Framework: Report from the EOSC Executive Board Working Groups FAIR and Architecture”, Directorate-General for Research and Innovation (European Commission), EOSC Executive Board, https://doi.org/10.2777/620649

Courtot, M, Gibson, F, Lister, AL, Malone, J, Schober, D, Brinkman, RR, Ruttenberg, A, 2009. ‘MIREOT: the Minimum Information to Reference an External Ontology Term‘, Nature Precedings. https://doi.org/10.1038/npre.2009.3574.1

Cox SJD, Gonzalez-Beltran AN, Magagna B, Marinescu M-C (2021) Ten simple rules for making a vocabulary FAIR. PLoS Comput Biol 17(6): e1009041. https://doi.org/10.1371/journal.pcbi.1009041

d’Aquin, M, Noy, NF, 2012. ‘Where to publish and find ontologies? A survey of ontology libraries‘, Journal of Web Semantics, vol. 11, pp. 96-111. https://doi.org/10.1016/j.websem.2011.08.005

Devaraju A., Huber R., 2021, An automated solution for measuring the progress toward FAIR research data, Patterns, Volume 2, Issue 11,https://doi.org/10.1016/j.patter.2021.100370.

Ding, L., T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. C. Doshi, and J. Sachs, "Swoogle: A Search and Metadata Engine for the Semantic Web", InProceedings, Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management , November 2004

Dutta, B, Toulet, A, Emonet, V, Jonquet, C, 2017. ‘New Generation Metadata Vocabulary for Ontology Description and Publication‘, in: Garoufallou, E, Virkus, S, Siatri, R, Koutsomiha, D (eds.), Metadata and Semantic Research, Communications in Computer and Information Science. Springer International Publishing, Cham, pp. 173-185. https://doi.org/10.1007/978-3-319-70863-8_17

European Commission Expert Group on FAIR Data, 2018. Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data. Directorate-General for Research and Innovation (European Commission), Brussels. https://doi.org/10.2777/1524




FAIRsFAIR “Fostering FAIR Data Practices In Europe” has received funding from the European Union’s Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Goldfarb, D, Le Franc, Y. 2017. 'Enhancing the Discoverability and Interoperability of Multi-Disciplinary Semantic Repositories', in: S4BioDiv@ISWC. Presented at the 2nd International Workshop on Semantics for Biodiversity co-located with 16th International Semantic Web Conference (ISWC 2017), Vienna.

Gruber, T.R., 1993. 'A Translation Approach to Portable Ontology Specifications', Knowledge Acquisition, vol. 5, pp. 199–220.

Guarino, N, Oberle, D, Staab, S, 2009. 'What Is an Ontology?', in: Staab, S, Studer, R (eds.), Handbook on Ontologies, International Handbooks on Information Systems. Springer, Berlin, Heidelberg, pp. 1–17. https://doi.org/10.1007/978-3-540-92673-3_0

Hartmann, J, Palma, R, Sure, Y, Suárez-Figueroa, MC, Haase, P, Gómez-Pérez, A, Studer, R, 2005. 'Ontology Metadata Vocabulary and Applications', in: Meersman, R, Tari, Z, Herrero, P (eds.), On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 906–915. https://doi.org/10.1007/11575863_112

European Commission, Directorate-General for Research and Innovation, Hellström, M., Heughebaert, A., Kotarski, R., et al., A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC), Publications Office, 2020, https://data.europa.eu/doi/10.2777/926037

Hugo, W, Le Franc, Y, Coen, G, Bonino, L, & Parland-von Essen, J,. (2020, October 30). Report on the FAIR Semantics workshop to discuss ontology design good practices and evaluate the first recommendations and the roadmap (M2.6) (Version 1.0). Zenodo. http://doi.org/10.5281/zenodo.431430

R. Jackson, N. Matentzoglu, J. A. Overton, R. Vita, J. P. Balhoff, P. L. Buttigieg, S. Carbon, M. Courtot, A. D. Diehl, D. M. Dooley, W. D. Duncan, N. L. Harris, M. A. Haendel, S. E. Lewis, D. A. Natale, D. Osumi-Sutherland, A. Ruttenberg, L. M Schriml, B. Smith, C. J. Stoeckert Jr., N. A. Vasilevsky, R. L. Walls, J. Zheng, C. J. Mungall, B. Peters, OBO Foundry in 2021: operationalizing open data principles to evaluate ontologies, Database, Volume 2021, 2021, baab069, https://doi.org/10.1093/database/baab069

Karray, M. , Otte, N. , Rai, R. , Ameri, F. , Kulvatunyou, B. , Smith, B. , Kiritsis, D. , Will, C. and Arista, R. (2021), The Industrial Ontologies Foundry (IOF) perspectives, Industrial Ontology Foundry (IOF) - achieving data interoperability Workshop, International Conference on Interoperability for Enterprise Systems and Applications, Tarbes, , [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=925879 (Accessed January 31, 2022)

Kulhanek, T. and Le Franc, Y., Service for Digital Annotation of Scientific Data, July 2019, Conference: IEEE EMBC 2019, At: Berlin, Germany

Kulvatunyou, B, Wallace, EK, Kiritsis, D, Smith, B, Will, C, 2018. The Industrial Ontologies Foundry Proof-of-Concept Project.



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Le Franc, Yann, Parland-von Essen, Jessica, Bonino, Luiz, Lehväslaiho, Heikki, Coen, Gerard, & Staiger, Christine. (2020). D2.2 FAIR Semantics: First recommendations (1.0 DRAFT). FAIRsFAIR. https://doi.org/10.5281/zenodo.3707985

Matentzoglu, N, Malone, J, Mungall, C, Stevens, R, 2018. 'MIRO: guidelines for minimum information for the reporting of an ontology.' Journal of Biomedical Semantics, vol. 9: 6. https://doi.org/10.1186/s13326-017-0172-7

Mazimwe, A., Hammouda, I. and Gidudu, A., 2021. Implementation of FAIR Principles for Ontologies in the Disaster Domain: A Systematic Literature Review. ISPRS International Journal of Geo-Information, 10(5), p.324.

Molloy, Laura, McQuilton, Peter, & Le Franc, Yann. (2021). EOSC Co-creation funded project 074: Delivery of a proof of concept for terms4FAIRskills: Technical report (1.0.0). Zenodo. https://doi.org/10.5281/zenodo.4772741

Musen, M.A. The Protégé project: A look back and a look forward. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. DOI: 10.1145/2557001.25757003.

Neuhaus, F, 2017. 'On the definition of 'ontology,'' in: Proceedings of the Joint Ontology Workshops 2017. Presented at the Episode 3: The Tyrolean Autumn of Ontology, Sun SITE Central Europe (CEUR), Bozen-Bolzano, Italy.

Obrst, L, 2010. 'Ontological Architectures', in: Poli, R, Healy, M, Kameas, A (eds.), Theory and Applications of Ontology: Computer Applications. Springer Netherlands, Dordrecht, pp. 27–66. https://doi.org/10.1007/978-90-481-8847-5_2

Obrst, L, 2003. 'Ontologies for semantically interoperable systems', in: CIKM '03: Proceedings of the Twelfth International Conference on Information and Knowledge Management. Presented at the The twelfth international conference on Information and knowledge management, New Orleans, Louisiana, USA, pp. 366-369. https://doi.org/10.1145/956863.956932

Ong, E, Xiang, Z, Zhao, B, Liu, Y, Lin, Y, Zheng, J, Mungall, C, Courtot, M, Ruttenberg, A, He, Y, 2017. 'Ontobee: A linked ontology data server to support ontology term dereferencing, linkage, query and integration', Nucleic Acids Research, vol. 45, D347-D352. https://doi.org/10.1093/nar/gkw918

Poveda-Villalón, María, Asunción Gómez-Pérez, and Mari Carmen Suárez-Figueroa. "OOPS!(Ontology Pitfall Scanner!): An on-line tool for ontology evaluation." International Journal on Semantic Web and Information Systems (IJSWIS) 10.2 (2014): 7-34.

M. Poveda-Villalón, P. Espinoza-Arias, D. Garijo, and O. Corcho, "Coming to Terms with FAIR Ontologies," in Knowledge Engineering and Knowledge Management, Cham, 2020, pp. 255–270, doi: 10.1007/978-3-030-61244-3_18

Rauber, A, Asmi, A, van Uytvanch, D, Proell, S, 2015. Data Citation of Evolving Data: Recommendations of the Working Group on Data Citation (WGDC). http://doi.org/10.15497/RDA00016



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

Schober, D, Smith, B, Lewis, SE, Kusnierczyk, W, Lomax, J, Mungall, C, Taylor, CF, Rocca-Serra, P, Sansone, S-A, 2009. 'Survey-based naming conventions for use in OBO Foundry ontology development', BMC Bioinformatics, vol. 10, 125. https://doi.org/10.1186/1471-2105-10-125

Seppälä, S, Ruttenberg, A, Smith, B, 2017. 'Guidelines for writing definitions in ontologies' Ciência da Informação, vol. 46 (1), pp. 73-88 Available at: http://revista.ibict.br/ciinf/article/view/4015

Smith, B, Ashburner, M, Rosse, C, Bard, J, Bug, W, Ceusters, W, Goldberg, L.J, Eilbeck, K, Ireland, A, Mungall, CJ, Leontis, N, Rocca-Serra, P, Ruttenberg, A, Sansone, S-A, Scheuermann, RH, Shah, N, Whetzel, PL, Lewis, S, 2007. 'The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration' Nature Biotechnology, vol. 25, pp. 1251-1255. https://doi.org/10.1038/nbt1346

Sterner, B., Upham, N., Gupta, P., Powell, C., & Franz, N. (2021). Wanted: Standards for FAIR taxonomic concept representations and relationships. Biodiversity Information Science and Standards, 5, e75587.

Vandenbussche, P-Y, Atemezing, GA, Poveda-Villalón, M, Vatant, B, 2017. 'Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web', Semantic Web, vol. 8, pp. 437-452. https://doi.org/10.3233/SW-160213

Weibel S., Kunze J., Lagoze C. et al. (1998) Dublin core metadata for resource discovery. Int. Eng. Task Force RFC, 2413, 132.

Whetzel, PL, Noy, NF, Shah, NH, Alexander, PR, Nyulas, C, Tudorache, T, Musen, MA, 2011. 'BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications.' Nucleic Acids Research, vol. 39, pp. W541-W545. https://doi.org/10.1093/nar/gkr469

Wilkinson, MD, Dumontier, M, Aalbersberg, IjJ, Appleton, G, Axton, M, Baak, A, Blomberg, N, Boiten, J-W, da Silva Santos, LB, Bourne, PE, Bouwman, J, Brookes, AJ, Clark, T, Crosas, M, Dillo, I, Dumon, O, Edmunds, S, Evelo, CT, Finkers, R, Gonzalez-Beltran, A, Gray, AJG, Groth, P, Goble, C, Grethe, JS, Heringa, J, 't Hoen, PAC, Hooft, R, Kuhn, T, Kok, R, Kok, J, Lusher, SJ, Martone, ME, Mons, A, Packer, AL, Persson, B, Rocca-Serra, P, Roos, M., van Schaik, R, Sansone, S-A, Schultes, E, Sengstag, T, Slater, T, Strawn, G, Swertz, MA, Thompson, M, van der Lei, J, van Mulligen, E, Velterop, J, Waagmeester, A, Wittenburg, P, Wolstencroft, K, Zhao, J, Mons, B, 2016. 'The FAIR Guiding Principles for scientific data management and stewardship.' Scientific Data 3, vol. 160018. https://doi.org/10.1038/sdata.2016.18

Wilkinson, MD, Dumontier, M, Sansone, S-A, Bonino da Silva Santos, LO, Prieto, M. Batista, D, McQuilton, P, Kuhn, T, Rocca-Serra, P, Crosas, M, Schultes, E, 2019. 'Evaluating FAIR maturity through a scalable, automated, community-governed framework ' Scientific Data 6, 174. https://doi.org/10.1038/s41597-019-0184-5



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

# 7. Appendix A: Minimum Metadata Voting Results

| Choices | | Mandatory | | Recommended | | Optional | | | |
|---|---|---|---|---|---|---|---|---|---|
| | total # of votes | # of votes | % of votes | # of votes | % of votes | # of votes | % of votes | **decision** | **consensus** |
| **Semantic Artefact (inherits from dcat:Resource and dcat:Dataset) -** https://w3id.org/mod#SemanticArtefact? | | | | | | | | | |
| dct:title | **23** | 21 | 91.30% | 1 | 4.35% | 1 | 4.35% | mandatory | 86.95% |
| dct:license | **28** | 24 | 85.71% | 3 | 10.71% | 1 | 3.57% | mandatory | 78.57% |
| dct:identifier | **28** | 22 | 78.57% | 3 | 10.71% | 3 | 10.71% | mandatory | 67.85% |
| dct:accessRights | **31** | 24 | 77.42% | 4 | 12.90% | 3 | 9.68% | mandatory | 66.12% |
| dct:creator | **31** | 24 | 77.42% | 5 | 16.13% | 2 | 6.45% | mandatory | 66.12% |
| dct:created | **13** | 9 | 69.23% | 2 | 15.38% | 2 | 15.38% | mandatory | 53.83% |
| dct:description | **28** | 19 | 67.86% | 6 | 21.43% | 3 | 10.71% | mandatory | 51.77% |
| dcat:contactPoint | **31** | 20 | 64.52% | 8 | 25.81% | 3 | 9.68% | mandatory | 46.76% |
| owl:versionIRI | **13** | 8 | 61.54% | 4 | 30.77% | 1 | 7.69% | mandatory | 42.29% |
| dct:modified | **28** | 16 | 57.14% | 9 | 32.14% | 3 | 10.71% | mandatory | 35.70% |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Choices | | Mandatory | | Recommended | | Optional | | | |
|---|---|---|---|---|---|---|---|---|---|
| dcat:keyword | **31** | 17 | 54.84% | 11 | 35.48% | 3 | 9.68% | mandatory | 32.24% |
| mod:acronym | **13** | 7 | 53.85% | 5 | 38.46% | 1 | 7.69% | mandatory | 30.75% |
| dcat:landingPage | **31** | 15 | 48.39% | 12 | 38.71% | 4 | 12.90% | mandatory | 22.56% |
| dct:publisher | **28** | 13 | 46.43% | 13 | 46.43% | 2 | 7.14% | recommended | 19.62% |
| dct:subject | **13** | 6 | 46.15% | 5 | 38.46% | 2 | 15.38% | mandatory | 19.21% |
| dct:type | **22** | 10 | 45.45% | 9 | 40.91% | 3 | 13.64% | mandatory | 18.16% |
| dct:issued | **27** | 12 | 44.44% | 11 | 40.74% | 4 | 14.81% | mandatory | 16.65% |
| dcat:theme | **30** | 12 | 40.00% | 10 | 33.33% | 8 | 26.67% | mandatory | 9.98% |
| dct:conformsTo | **30** | 12 | 40.00% | 13 | 43.33% | 5 | 16.67% | recommended | 14.98% |
| dct:language | **28** | 11 | 39.29% | 13 | 46.43% | 4 | 14.29% | recommended | 19.62% |
| mod:URI | **13** | 5 | 38.46% | 2 | 15.38% | 6 | 46.15% | optional | 19.21% |
| dcat:distribution | **25** | 8 | 32.00% | 11 | 44.00% | 6 | 24.00% | recommended | 15.98% |
| dct:contributor | **13** | 4 | 30.77% | 7 | 53.85% | 2 | 15.38% | recommended | 30.75% |
| dct:rights | **23** | 7 | 30.43% | 13 | 56.52% | 3 | 13.04% | recommended | 34.77% |
| dct:temporal | **24** | 6 | 25.00% | 9 | 37.50% | 9 | 37.50% | recommended | 6.23% |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Choices | | Mandatory | | Recommended | | Optional | | | |
|---|---|---|---|---|---|---|---|---|---|
| dcat:qualifiedRelation | **29** | 7 | 24.14% | 7 | 24.14% | 15 | 51.72% | optional | 27.57% |
| mod:status | **13** | 3 | 23.08% | 8 | 61.54% | 2 | 15.38% | recommended | 42.29% |
| odrl:hasPolicy | **23** | 5 | 21.74% | 7 | 30.43% | 11 | 47.83% | optional | 21.72% |
| prov:qualifiedAttribution | **23** | 5 | 21.74% | 4 | 17.39% | 14 | 60.87% | optional | 41.29% |
| prov:wasGeneratedBy | **26** | 5 | 19.23% | 10 | 38.46% | 11 | 42.31% | optional | 13.44% |
| dct:relation | **22** | 4 | 18.18% | 8 | 36.36% | 10 | 45.45% | optional | 18.16% |
| dct:isReferencedBy | **28** | 5 | 17.86% | 11 | 39.29% | 12 | 42.86% | optional | 14.26% |
| schema:includedInDataCatalog | **12** | 2 | 16.67% | 3 | 25.00% | 7 | 58.33% | optional | 37.48% |
| mod:competencyQuestion | **13** | 2 | 15.38% | 1 | 7.69% | 10 | 76.92% | optional | 65.38% |
| dct:accrualPeriodicity | **25** | 3 | 12.00% | 10 | 40.00% | 12 | 48.00% | optional | 21.98% |
| dct:spatial | **24** | 2 | 8.33% | 11 | 45.83% | 11 | 45.83% | optional | 18.73% |
| mod:usedEngineeringMethodology | **12** | 1 | 8.33% | 6 | 50.00% | 5 | 41.67% | recommended | 24.98% |
| dcat:temporalResolution | **25** | 2 | 8.00% | 8 | 32.00% | 15 | 60.00% | optional | 39.98% |
| mod:hasFormalityLevel | **13** | 1 | 7.69% | 6 | 46.15% | 6 | 46.15% | recommended | 19.21% |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

| Choices | | Mandatory | | Recommended | | Optional | | | |
|---|---|---|---|---|---|---|---|---|---|
| dcat:spatialResolutionInMeters | **25** | 1 | 4.00% | 6 | 24.00% | 18 | 72.00% | optional | 57.99% |
| dct:accrualMethod | **13** | 0 | 0.00% | 8 | 61.54% | 5 | 38.46% | recommended | 42.29% |
| **Semantic Artefact Distribution (inherits from dcat:Distribution)** - https://w3id.org/mod#SemanticArtefactDistribution | | | | | | | | | |
| dcat:mediaType | **17** | 11 | 64.71% | 5 | 29.41% | 1 | 5.88% | mandatory | 47.05% |
| dct:format | **17** | 11 | 64.71% | 5 | 29.41% | 1 | 5.88% | mandatory | 47.05% |
| dct:title | **15** | 9 | 60.00% | 3 | 20.00% | 3 | 20.00% | mandatory | 39.98% |
| dcat:accessURL | **17** | 10 | 58.82% | 4 | 23.53% | 3 | 17.65% | mandatory | 38.22% |
| mod:hasRepresentationLanguage | **10** | 5 | 50.00% | 2 | 20.00% | 3 | 30.00% | mandatory | 24.98% |
| mod:hasSyntax | **10** | 5 | 50.00% | 2 | 20.00% | 3 | 30.00% | mandatory | 24.98% |
| dct:accessRights | **17** | 8 | 47.06% | 6 | 35.29% | 3 | 17.65% | mandatory | 20.57% |
| dcat:downloadURL | **17** | 7 | 41.18% | 8 | 47.06% | 2 | 11.76% | recommended | 20.57% |
| dct:rights | **17** | 6 | 35.29% | 8 | 47.06% | 3 | 17.65% | recommended | 20.57% |
| dct:description | **17** | 5 | 29.41% | 8 | 47.06% | 4 | 23.53% | recommended | 20.57% |
| dct:issued | **17** | 5 | 29.41% | 6 | 35.29% | 6 | 35.29% | recommended | 2.92% |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558

FAIRSFAIR
Fostering Fair Data Practices in Europe

| Choices | | Mandatory | | Recommended | | Optional | | | |
|---|---|---|---|---|---|---|---|---|---|
| dct:modified | **16** | 4 | 25.00% | 9 | 56.25% | 3 | 18.75% | recommended | 34.36% |
| mod:definitionProperty | **10** | 2 | 20.00% | 6 | 60.00% | 2 | 20.00% | recommended | 39.98% |
| dcat:accessService | **16** | 3 | 18.75% | 8 | 50.00% | 5 | 31.25% | recommended | 24.98% |
| dcat:packageFormat | **17** | 2 | 11.76% | 7 | 41.18% | 8 | 47.06% | optional | 20.57% |
| dct:conformsTo | **17** | 2 | 11.76% | 9 | 52.94% | 6 | 35.29% | recommended | 29.39% |
| mod:usedEngineeringTool | **10** | 1 | 10.00% | 4 | 40.00% | 5 | 50.00% | optional | 24.98% |
| mod:prefLabelProperty | **10** | 1 | 10.00% | 7 | 70.00% | 2 | 20.00% | recommended | 54.99% |
| mod:synonymProperty | **10** | 1 | 10.00% | 5 | 50.00% | 4 | 40.00% | recommended | 24.98% |
| odrl:hasPolicy | **15** | 1 | 6.67% | 9 | 60.00% | 5 | 33.33% | recommended | 39.98% |
| dcat:compressFormat | **17** | 1 | 5.88% | 4 | 23.53% | 12 | 70.59% | optional | 55.87% |
| dcat:temporalResolution | **17** | 1 | 5.88% | 3 | 17.65% | 13 | 76.47% | optional | 64.70% |
| dcat:byteSize | **17** | 0 | 0.00% | 5 | 29.41% | 12 | 70.59% | optional | 55.87% |
| dcat:spatialResolutionInMeters | **17** | 0 | 0.00% | 2 | 11.76% | 15 | 88.24% | optional | 82.35% |



FAIRsFAIR "Fostering FAIR Data Practices In Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 grant agreement 831558