

@ontocommons



linkedin.com/company/ontocommons



info@ontocommons.eu



Towards Materials and Manufacturing Commons - the enablers Digital Marketplaces, FAIR Principles and Ontologies

Berlin – April 4th – 6th 2023

FAIR Semantics

A requirement for interoperability

Yann Le Franc – CEO and Sicentific Director – e-Science Data Factory (France)

ylefranc@esciencefactory.com

ORCID: 0000-0003-4631-418X



ONTO COMMONS Why FAIR Semantics?

- Semantic landscape heterogeneous:
 - ✓ Different levels of semantics (from weak to strong)
 - ✓ Different formats: RDF/XML, OWL, Turtle,...
 - ✓ Different names: ontologies, controlled vocabularies, thesauri, codelists, ...
 - ✓ Different community's best practices
 - ✓ Different ways of sharing/publishing

Hard to find and reuse ontologies within and across domains + Interoperability problems = NOT FAIR

ONTO COMMONS Why FAIR Semantics?

- Semantic landscape heterogeneous:
 - ✓ Different levels of semantics (from weak to strong)
 - ✓ Different formats: RDF/XML, OWL, Turtle,...
 - ✓ Different names: ontologies, controlled vocabularies, thesauri, codelists, ...
 - ✓ Different community's best practices
 - ✓ Different ways of sharing/publishing

Hard to find and reuse ontologies within and across domains + Interoperability problems = NOT FAIR

FAIR Principle I2. (meta)data use vocabularies that follow FAIR principles



COMMONS What do we mean by FAIR Semantics?

FAIR Semantics means that semantic artefacts should adhere to the FAIR principles. We consider semantic artefacts as a specific type of data, used to describe or annotate other data, i.e. as metadata.



ONTO COMMONS What do we mean by FAIR Semantics?

FAIR Semantics means that **semantic artefacts** should adhere to the FAIR principles. We consider semantic artefacts as a specific type of data, used to describe or annotate other data, i.e. as metadata.

A *semantic artefact* is defined within our 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.



ONTO MONS What does it mean for an ontology to be FAIR?



February 25, 2022

- Recommendation aligned with RFC 2119 (MUST, SHOULD, MAY)
 - 9 MUST
 - 7 SHOULD
 - **○** 1 MAY
 - 1 Undetermined

D2.8 FAIR Semantics Recommendations Third Iteration

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.

(b) Yann Le Franc; (b) Luiz Bonino; (b) Hanna Koivula; (b) Jessica Parland-von Essen; (b) Robert Pergl



207 175 downloads See more details. Indexed in **OpenAIRE**



https://github.com/FAIRsFAIR/FAIRSemantics



How did we built them?





FAIRsFAIR workshop co-located with RDA P14

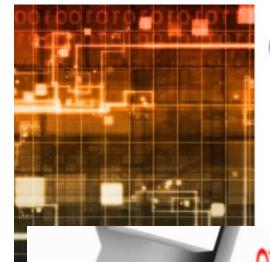
AND FAIR



Recommendations for FAIR Semantics

Online WORKSHOP

15 October 2020 13:00-17:00





Workshop Minimum metadata schema for semantic artefacts





4 June 2021

09:30 - 13:00 CEST



ONTO COMMONS Collaboration with RDA









FAIR Semantics, the Semantic Web Universe and Everything

Home » Plenaries » FAIR Semantics, the Semantic Web Universe and Everything



By Yann Le Franc

Group(s) submitting the application: Vocabulary Services IG

Meeting objectives:

The VSSIG is a unique platform for international Semantic Web experts to address key issues in this field and to share tips and tricks as well as ongoing activities. One of the hot discussion topics within the group is about FAIR Semantics—that is, how to make semantic artifacts (ontologies, controlled vocabularies, thesauri, glossaries...) understandable by



ONTO COMMONS Collaboration with RDA







- RDA VSSIG Task Group on Minimum Metadata C. Jonquet & L. Bonino
 - Define a minimum metadata schema for FAIR semantic artefacts (i.e. ontologies,...
 - Define a DCAT profile to publish collections of semantic artefacts
- RDA VSSIG Task Group on FAIR Semantic repositories A. Kokkinaki & G. Coen
 - Evaluate recommendations from the perspective of ontology repositories
 - -Establish a list of possible technical implementation of the recommendations



ONTO STATE COMMONS Not reinventing the wheel



- FAIR Semantics recommendations are linked to similar work.
 - Best practices for implementing fair vocabularies and ontologies on the web - Daniel Garijo and Maria Poveda (2020)
 - ◆ Ten simple rules for making a vocabulary FAIR Cox et al. (2021)
 - OBO Foundry principles
 - IOF principles



ONTO COMMONS Applying FAIR to Semantic Artefacts COMMONS

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.



ONTO MADDINISTRIAN PRINCIPINA Are ontologies for industry FAIR?

- What are the existing ontologies or vocabularies relevant for industry?
- What are their characteristics (format, logic, serialisations, ...)?
- Are they aligned with Top Level Ontologies?

Landscape analysis of domain ontologies for industry



ONTO COMMONS How did we collect our sample?



OntoCommons Ontology Catalogue Home /

https://ontocommons.eu/node/146



- Internal project knowledge
- Workshops with experts
- Survey

Metadata for OntoCommons ontology catalogue

This survey is oriented to people who are aware of any ontology that could be useful for materials, manufacturing or related domains.

Our final goal is to develop an ontology catalogue in order to provide the materials and manufacturing communities with the most suitable ontologies in this area. In addition the catalogue also contains general domain ontologies frequently used across domains

The estimated time required to complete the questionnaire is of 10 minutes. Once the form about an ontology is submitted it will be manually assessed and automatically processed. After this, the ontology will be included in the future Onto Commons Catalogue. Please note that there is a manual component in the process, therefore the on-line catalogue will not be updated immediately after the submission.

The questionnaire does not include any personal guestion and the confidentiality of the answers will be preserved. We only ask for an email address just in case you want to obtain information about the results we produce.

This questionnaire is being performed in the context of the OntoCommons HORIZON2020 project

If you have any question or comment about the questionnaire contact ontocommons.registry[at]delicias.dia.fi.upm.es

13.



ONTO COMMONS Sharing our dataset COMMONS Sharing our dataset

Home SPARIOL

https://data.ontocommons.linkeddata.es/index

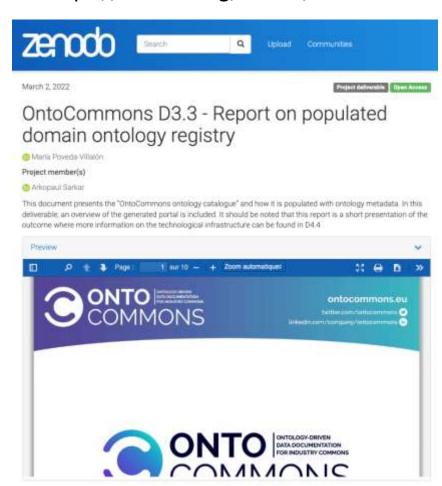
OntoCommons ontology catalogue

On the Semantic Web, ontologies define the concepts and relationships used to describe a given domain and annotate data about it. In the OntoCommons Horizon CSA we are collecting ontologies about materials, construction, manufacturing and other industries. Here you can find the list of ontologies we have identified so far. You can also propose ontologies to be included in the catalogue by filling in the form.

Ontology catalogue overview

Ontology	URI	Licensed?	Ontology Language	Syntax	Domain	Natural Language
Battery (NterFace Ontology (BattINFO)	redology - CCC LO Universal DML Butter Electrochemistry Electrochemistry Electrochemistry					
Battery Value Chain Ontology (BVCC)	*	CC-NY	OWL	Turtio	BatteryNatusChate MiningOfflameryMaterials TelliningOfflatteryMaterials BatteryMaterials BetteryRespond	-
Building antology	*	OC-RY	OWL	Tartis	Construction Renewation	eng
CIF-Ontology	*	OC-NY	OWL	Tlertin	ManufaleScience Chemistry Physics Crystallography	emp
Collaborative Manufacturing Services Ontology	*	MIT	CIWL	RDF/XML	ManufacturingAndSupplyChainDumaina	ang
Crystallography Domain Ontology	*	CC-RA	CIWS.	Turthi	Materialitisience Crystofography	-
Digital Construction Energy Systems	*	CC-RY	OWL	Yarna	BigftaConstruction	ang
Digital Construction Entities	*	OC-RY	CIWIL	Turthis	BigitatConstruction	MITTER
Digital Construction Materials	*	CC-RY	GWL.	Tlartin	DigitalConstruction	arro

https://zenodo.org/record/6504584



14.



Our dataset

- 130 ontologies
- Classified in 5 domains
 - Physics and Chemistry
 - Mechanical and Industrial Engineering
 - Thermal and Process Engineering
 - Material Sciences and Engineering
 - Computer Sciences, Systems and Electrical Engineering





OntoCommons | pre-printed version

Report on selsting domain ontologies in

Project Title	ect Title Ontology-driven data documentation for Industry Commons	
Project Acronym	OntoCommons	
Project Number	958371	
Type of project	CSA - Coordination and support action	
Topics	DT-NM8P-39-2020 - Towards Standardised Documentation of Data through taxonomies and ontologies (CSA)	
Starting date of Project	01 November 2020	
Duration of the project 36 months		
Nebsite www.ontocommons.eu		

Report on existing domain ontologies in identified domains

Work Package	Industrial Domain Ontologies	
Task	Domain-specific semantic Landscape Analysis	
Lead author	Yann Le Franc (eSDF)	
Contributors	Gerhard Goldbeck (GCL), Arkopaul Sarkar (ENIT), Jesper Friis (SINTEF), Maria Poveda Villalon (UPM), Alba Fernández Izquierdo (UPM), Hedi Karray (ENIT), Emna Amdouni (ENIT), Emilio Sanfilippo (CNR)	
Peer reviewers	Dimitris Kiritsis (UiO), John Breslin (NUIG)	
Version	Final	
Date	09/03/2022	

https://zenodo.org/record/6504553

28/06/2023 Towards Materials an



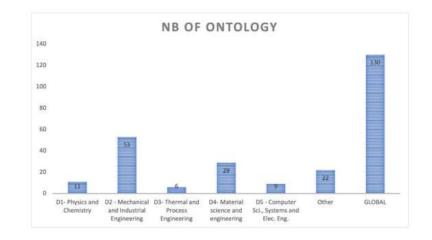


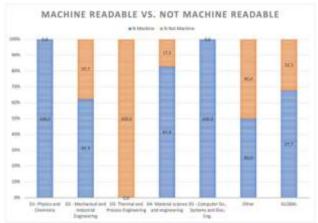


Project Title	Ontology-driven data documentation for Industry Commons		
Project Acronym	OntoCommons		
Project Number	958371		
Type of project	CSA - Coordination and support action		
Topics	DT-NM8P-39-2020 - Towards Standardised Documentation of Data through taxonomies and ontologies (CSA)		
Starting date of Project	01 November 2020		
Duration of the project	36 months		
Website	www.ontocommons.eu		

Report on existing domain ontologies in identified domains

Work Package	Industrial Domain Ontologies	
Task	Domain-specific semantic Landscape Analysis	
Lead author	Yann Le Franc (eSDF)	
Contributors	Gerhard Goldbeck (GCL), Arkopaul Sarkar (ENIT), Jesper Friis (SINTEF), Maria Poveda Villalon (UPM), Alba Fernández Izquierdo (UPM), Hedi Karray (ENIT), Emna Amdouni (ENIT), Emilio Sanfilippo (CNR)	
Peer reviewers	Dimitris Kiritsis (UiO), John Breslin (NUIG)	
Version	Final	
Date	09/03/2022	







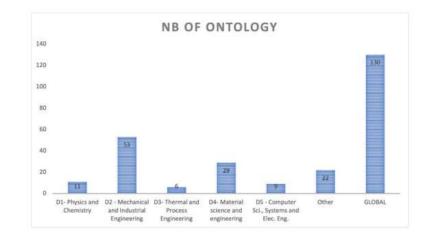


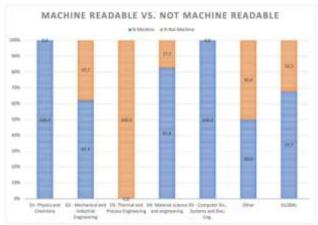


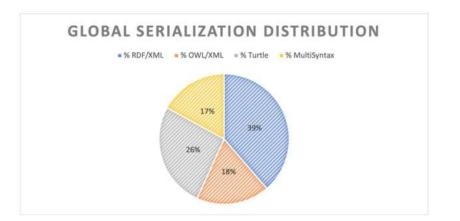
Project Title	Ontology-driven data documentation for Industry Commons		
Project Acronym	OntoCommons		
Project Number	958371		
Type of project	CSA - Coordination and support action		
Topics	DT-NM8P-39-2020 - Towards Standardised Documentation of Data through taxonomies and ontologies (CSA)		
Starting date of Project	01 November 2020		
Duration of the project	36 months		
Website	www.ontocommons.eu		

Report on existing domain ontologies in identified domains

Work Package	Industrial Domain Ontologies	
Task	Domain-specific semantic Landscape Analysis	
Lead author	Yann Le Franc (eSDF)	
Contributors Gerhard Goldbeck (GCL), Arkopaul Sarkar (ENIT), (SINTEF), Maria Poveda Villaion (UPM), Alba Fernánde (UPM), Hedi Karray (ENIT), Emna Amdouni (EN Sanfilippo (CNR)		
Peer reviewers	ewers Dimitris Kiritsis (UiO), John Breslin (NUIG)	
Version	Final	
Date	09/03/2022	









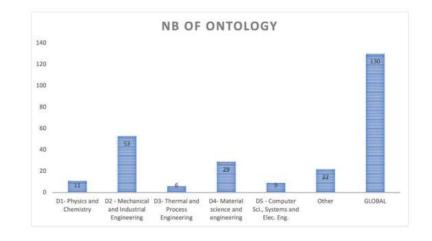


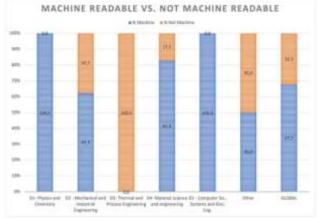


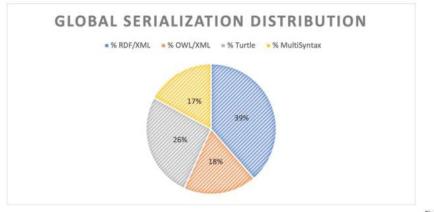
Project Title	Ontology-driven data documentation for Industry Commons		
Project Acronym	OntoCommons		
Project Number	958371		
Type of project	CSA - Coordination and support action		
Topics	DT-NM8P-39-2020 - Towards Standardised Documentation of Data through taxonomies and ontologies (CSA)		
Starting date of Project	01 November 2020		
Duration of the project	36 months		
Website	www.ontocommons.eu		

Report on existing domain ontologies in identified domains

Work Package	Industrial Domain Ontologies	
Task	Domain-specific semantic Landscape Analysis	
Lead author	Yann Le Franc (eSDF)	
Contributors	Gerhard Goldbeck (GCL), Arkopaul Sarkar (ENIT), Jesper Friis (SINTEF), Maria Poveda Villalon (UPM), Alba Fernández Izquierdo (UPM), Hedi Karray (ENIT), Emna Amdouni (ENIT), Emilio Sanfilippo (CNR)	
Peer reviewers	Dimitris Kiritsis (UiO), John Breslin (NUIG)	
Version	ion Final	
Date	09/03/2022	







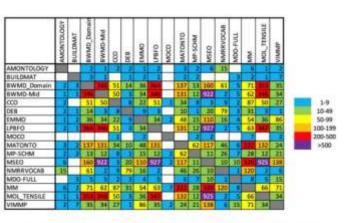
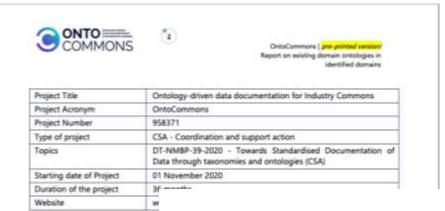
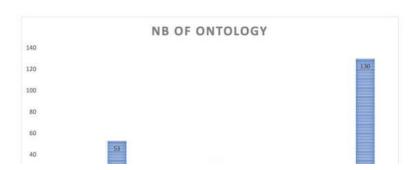


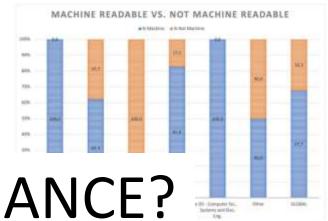
Figure 32 - Ontology overlap expressed as number of mappings between every pair of antologies from MatPortal.









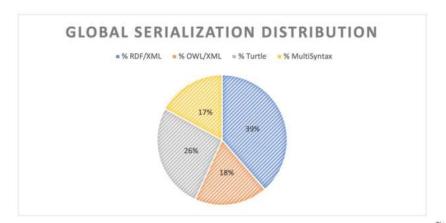


WHAT ABOUT FAIR COMPLIANCE?

Report on existin

domains

Work Package	Industrial Domain Ontologies	
Task	Domain-specific semantic Landscape Analysis	
Lead author	Yann Le Franc (eSDF)	
Contributors Gerhard Goldbeck (GCL), Arkopaul Sarkar (ENIT), Jes (SINTEF), Maria Poveda Villalon (UPM), Alba Fernández I (UPM), Hedi Karray (ENIT), Emna Amdouni (ENIT Sanfilippo (CNR)		
Peer reviewers	Dimitris Kiritsis (UiO), John Breslin (NUIG)	
Version	Final	
Date	09/03/2022	



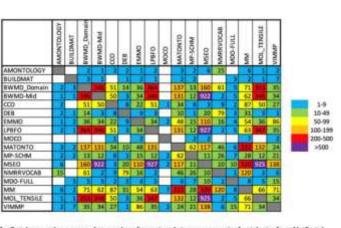


Figure 32 - Ontology overlap expressed as number of mappings between every pair of ontologies from MatPortal.



ONTO COMMONS Identify the recommendations relevant to ontologies

Rec#		Recommendation	Target
P-Rec	1	Globally Unique, Persistent and Resolvable Identifiers MUST be used for Semantic Artefacts, their content (terms/concepts/classes and relations), and their version	Ontology
P-Rec	2	Globally Unique, Persistent and Resolvable Identifiers MUST be used for Semantic Artefacts metadata records. Metadata and data must be published separately, even if it is managed jointly	Ontology/Repository
P-Rec	3	A common minimum metadata schema MUST be used to describe semantic artefacts and their content	Ontology
P-Rec	5	Semantic repositories MUST offer access to Semantic Artefacts and their content using community standard APIs and serializations to support both use/reuse and indexation by search engine	Repository
P-Rec	7	Repositories MUST offer a secure access protocol, and appropriate user access control functionalities	Repository
P-Rec	8	Human and machine-readable persistence policies for semantic artefacts metadata and data MUST be published	Repository
P-Rec	9	Semantic artefacts MUST be made available as a minimum portfolio of common serialization formats	Ontology/Repository
		The Semantic Artefact MUST be clearly licenced for use by machines and humans	Ontology
P-Rec	17	Provenance MUST be clear for both humans and machines	Ontology

Rec#		Recommendation	Target
P-Rec	4	Semantic artefacts and its content SHOULD be published in a trustworthy semantic repository	Ontology
P-Rec	11	A standardized knowledge representation language SHOULD be used for describing semantic artefacts	Ontology
P-Rec	12	Semantic mappings between the different elements of semantic artefacts SHOULD be published in machine readable format	Semantic Community
P-Rec	13	Crosswalks, mappings and bridging between semantic artefacts SHOULD be documented, published, and curated	Semantic Community
P-Rec	14	Standard vocabularies SHOULD be used to describe semantic artefacts	Ontology
P-Rec	15	Provenance information regarding the reuse of components from third-party semantic artefacts SHOULD be made explicit	Ontology
		Foundational Ontologies MAY be used to align semantic artefacts	Ontology
P-Rec	6	Build semantic artefact search engines that operate across different semantic repositories	

20.



Define the questions to answer for each of the relevant recommendations

- **○**—13 yes/no questions
- 1 or more questions
 for each recommendations

Rec#	Topic	Question
P-Rec 1	GUPRI	Does the SA have a persistent identifier of type purl, w3id or handle except for DOI?
P-Rec 1	GUPRI	Does the identifier resolve to a machine-readable format?
P-Rec 1	GUPRI	Does the SA provide a GUPRI for version?
P-Rec 3	Metadata	Does the SA have descriptive metadata?
P-Rec 14	Standard Vocabularies	Does SA's metadata use widely used vocabularies (dc, dct,)?
P-Rec 17	Provenance	Does the SA have provenance information?
P-Rec 17	Provenance	Does the SA use W3C Prov?
P-Rec 15	Provenance	Does the SA describe imports with provenance?
P-Rec 4	Publication	Is the SA published on a dedicated trusted semantic repository?
P-Rec 16	Licence	Does the SA have a license?
P-Rec 16	Licence	Is the license machine-readable?
P-Rec 11	Language	Does the SA use a standard knowledge representation such as SKOS, OWL?
P-Rec 10	TLO	Does the SA align with a Top-Level Ontology?

8/06/2023 Towards Materials an



ONTO COMMONS Establish a measurement method

Method

- If a recommendation is fulfilled we score 1 else we score 0
- If a recommendation is represented by several questions and if all these questions are answered positively we score 1 else we score 0

FAIR Score

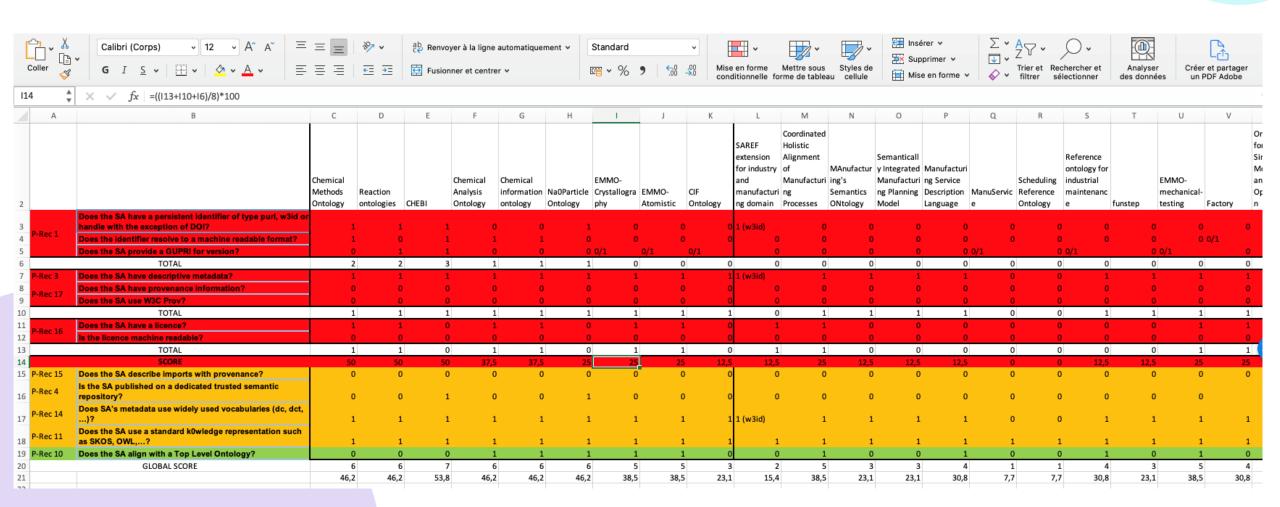
Percentage of mandatory recommendations fulfilled

Global FAIR Score

Percentage of all the fulfilled recommendations



Measurement in practice



28/06/2023 Towards Materials an



FAIRness assessment performed on 44 out of 74 machine-readable ontologies. Average FAIRness score by domain is

Domain	FAIR Score	Global FAIR Score
Physics and Chemistry	34,7 % (± 13,7 %)	42,7 % (± 8,7 %)
Mechanical and Industrial Engineering	18,8 % (± 14,4%)	27,8 % (± 11,8%)
Material Science and Engineering	28,8 % (± 21,3 %)	40,8 % (± 16,2 %)
Computer Science, Systems and Electrical Engineering	25 % (±19,1 %)	30,8 % (16,6)

- Physics and Chemistry is the domain with the highest FAIR Score on average.
- Allotrope ontology (in material science domain) being most FAIR ontology.
- no ontologies passed the threshold of minimally FAIR.

24.



- Great tool to get started (tested during Agrohackathon 08/22)
- Incomplete approach as several FAIR principles are not covered
- Based on high level recommendations: need to define practical implementations
- No consideration about the metadata content describing ontologies
- Comparaison with FOOPS!

25.



info@ontocommons.eu

Thank you for your attention





Join our community

Follow us on Twitter:

@ontocommons

Follow us on LinkedIn

linkedin.com/company/ontocommons

Subscribe to our Newsletter:

ontocommons.eu/newsletter





Identifier



Rec#	Recommendation	FAIR Principle
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
P-Rec. 2	Globally Unique, Persistent and Resolvable Identifiers MUST be used for Semantic Artefact Metadata Record. Metadata and data must be published separately, even if it is managed jointly	F1, F3





Metadata



Rec#	Recommendations	FAIR principle
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
P-Rec. 8	Human and machine-readable persistence policies for semantic artefacts metadata and data MUST be published	A2
P-Rec. 9	Semantic artefacts MUST be made available as a minimum portfolio of common serialization formats	I1
P-Rec. 14	Standard vocabularies SHOULD be used to describe semantic artefacts	12
P-Rec. 15	Provenance information regarding the reuse of components from third-party semantic artefacts SHOULD be made explicit	13, R1.2
P-Rec. 16	The semantic artefact MUST be clearly licenced for use by machines and humans	R1.1
P-Rec. 17	Provenance MUST be clear for both humans and machine	R1.2



« Semantic alignment »



Rec#	Recommendations	FAIR Principles
P-Rec. 10	Foundational Ontologies MAY be used to align semantic artefacts	11, 12, 13
P-Rec. 11	A standardized knowledge representation language SHOULD be used for describing complex logical relations (semantic artefact)	I1
P-Rec. 12	Semantic mappings between the different elements of semantic artefacts SHOULD be published in machine-readable formats	I1, I3, R1.3
P-Rec. 13	Crosswalks, mappings and bridging between semantic artefacts SHOULD be documented, published and curated	R1.2, R1.3





Repository



Rec#	Recommendation	FAIR Principles
P-Rec. 4	Semantic Artefact and its content SHOULD be published in a trustworthy semantic repository	F4
P-Rec.5	Semantic repositories MUST offer access to Semantic Artefacts and their content using community standard APIs and serializations to support both use/reuse and indexation by search engines	F4, A1, A1.1
P- Rec. 6	Build semantic artefacts' search engines that operate across different semantic repositories	F4
P-Rec. 7	Repositories MUST offer a secure access protocol and appropriate user access control functionalities	A1.2