FAIRness assessment for semantic artefacts

Clement Jonquet (INRAE, U. Montpellier)

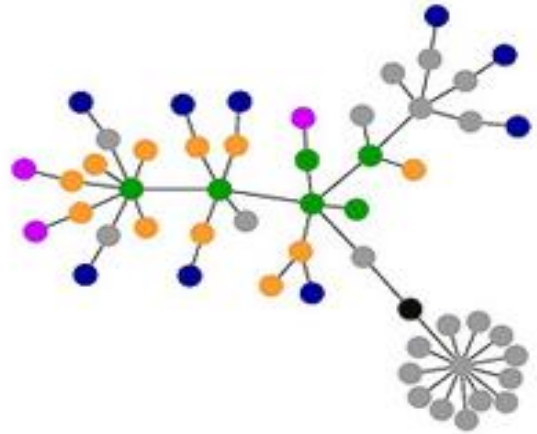
Session on
FAIR, Data Quality Management and Research Metrics Assessment

Funded by the European Union

Semantic Artefacts help to make data FAIR and have themselves to be FAIR

I2. (Meta)data use vocabularies that follow FAIR principl

How to go from ”principles” to specialized criteria to measure to which level ontologies respect the FAIR Principles?



From I1, I2 and I3, ontologies/vocabularies are a key element to achieve the FAIR Principles

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Research data
» Publication characteristics

Comment: The FAIR Gu… Principles for scient… data management and stewardship

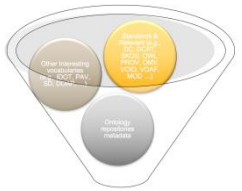
Mark D. Wilkinson *et al.*#

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

- As any data, semantic artefacts need to be FAIR

https://doi.org/10.7490/f1000research.1115343.1



- Assessing the Practice of Ontology Metadata: A Survey

https://hal-lirmm.ccsd.cnrs.fr/lirmm-02315001

- MOD1.4: A metadata vocabulary for ontology description and publication

https://dx.doi.org/10.1007/978-3-319-70863-8_17

- Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal

https://doi.org/10.1007/s13740-018-0091-5

- FAIRsFAIR minimum metadata profile for semantic artefacts

https://hal.science/hal-04106533

FAIR-IMPACT tasks related to this subject

- T4.2 Semantic artefact lifecycle and their catalogues

…establish guidelines and community practices with respect to the lifecycle of semantic artefacts from creation to sharing and reuse via catalogues.

  - T4.2.1 FAIR semantic artefact by design
  - T4.2.2 Interoperable semantic artefact catalogues
  - T4.2.3 Standardized semantic artefact metadata and catalogues APIs

- T5.3 Semantic artefact assessment methodology
  (see recent deliverable https://zenodo.org/record/8305173)

- M. Poveda, D. Garijo, Y. Le Franc, A. Gonzalez-Beltran, C. Jonquet … and more…

We now have an ontology/semantic artefact development methodology integrating FAIRness assessment

zenodo Search Upload Communities

Zenodo.org will be unavailable for 2 hours on September 29th from 06:00-08:00 UTC. See announcement.

August 31, 2023 Project milestone Open Access

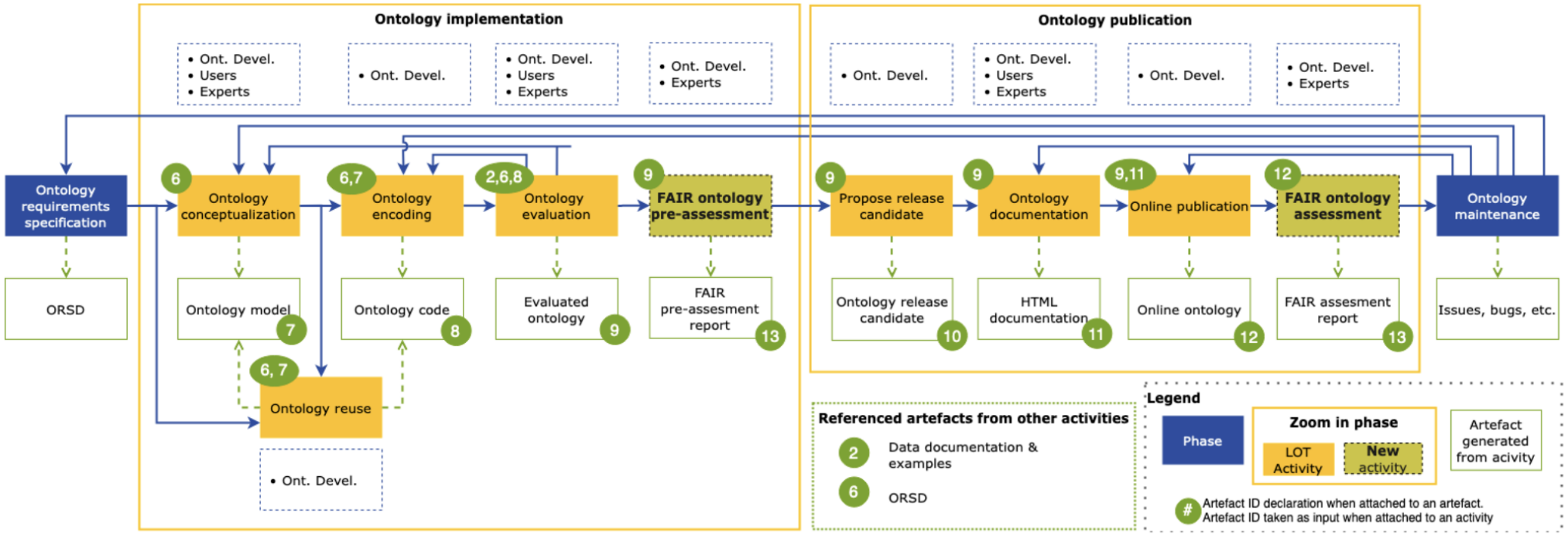
M5.3 Semantic artefact assessment methodology

Garijo, Daniel; Poveda-Villalón, María; Flohr, Pascal; Gonz
Maaike

Semantic artefacts (i.e., ontologies, vocabularies and SKOS taxonom
construction of, and help validate many existing Knowledge Graphs.
proposed (Poveda-Villalón et al. 2020; Garijo and Poveda-Villalón 20
2023) to align semantic artefact best practices against the Findable,
(FAIR principles) (Wilkinson et al. 2016). Based on these guidelines, r
(Garijo et al. 2021; Amdouni et al. 2022a; 2022b) in order to guide us
FAIR principles. However, different tests are based on different interp
scores and checks for semantic artefacts. To the best of our knowle
types of tests to perform in semantic artefacts, in order to map exist

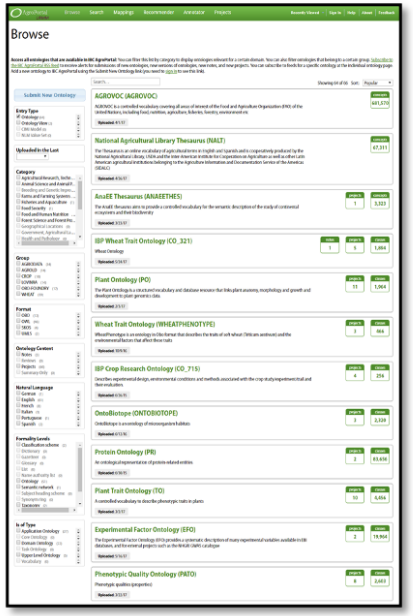
In this document, we propose such a methodology. We do so by taki
semantic artefacts into smaller parts (their code, content, ontology r
different stages of their development process. We build on the Linke
al. 2022), adding a "FAIR assessment" module, and, for each activity,
semantic artefact FAIR assessment validators: FOOPS! (Garijo et al.

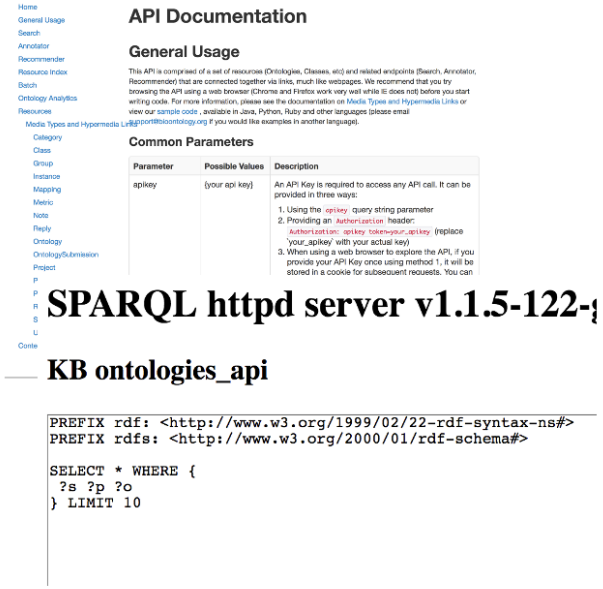
The rest of the document outlines our methodology, describes each
and guidelines.

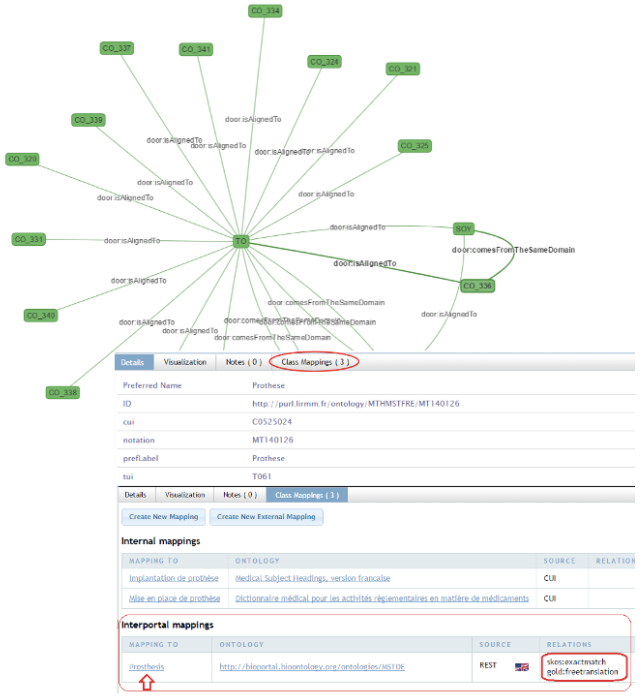


Semantic Artefact Catalogues (aka. ontology repositories, terminology services…) are necessary for FAIR

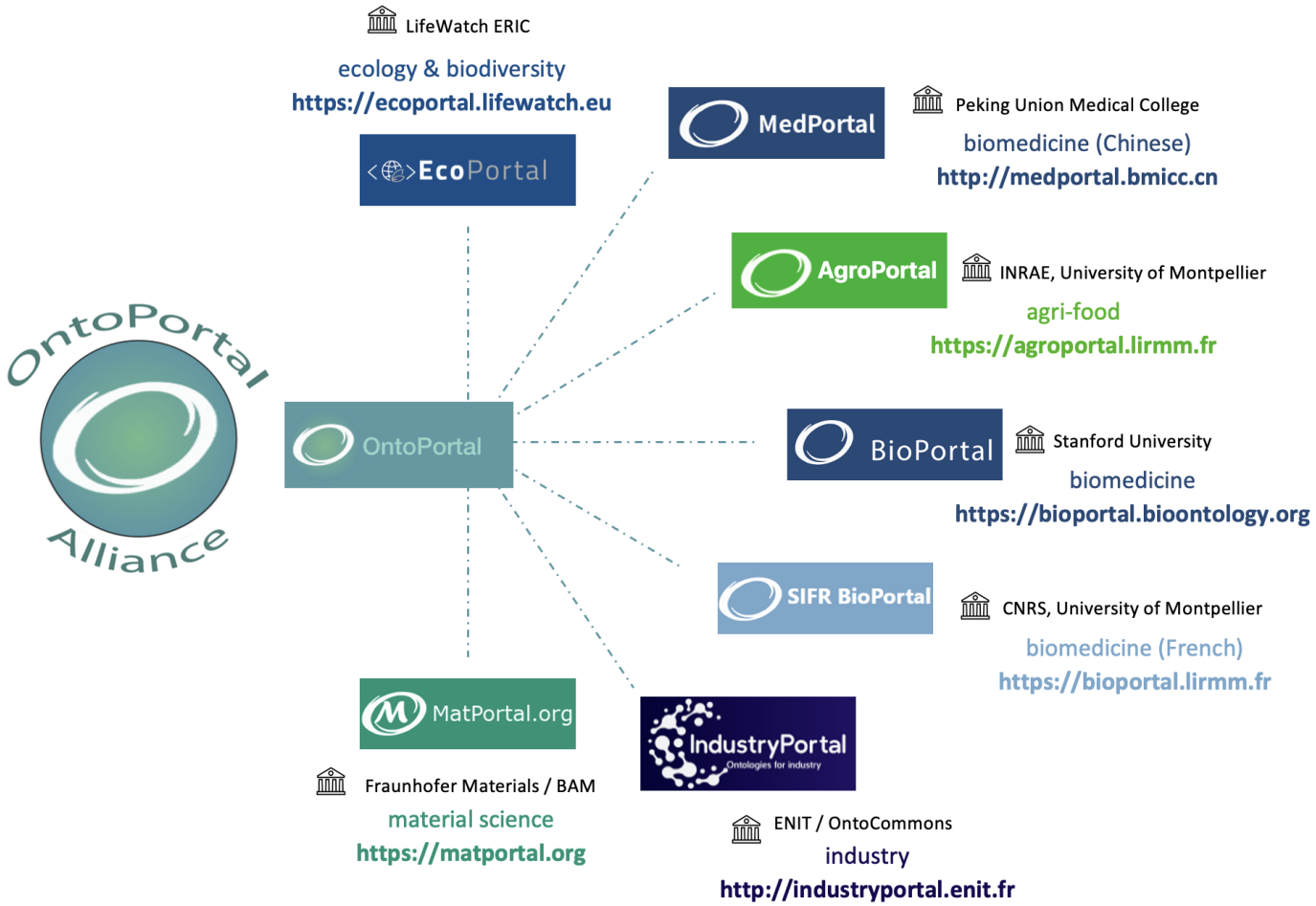
Findable **A**ccessible **I**nteroperable **R**e-usable







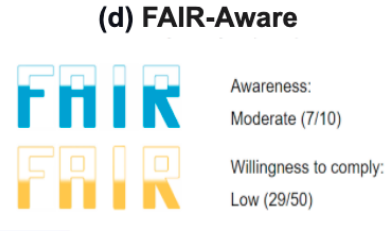
OntoPortal Alliance: Synchronizing and mutualizing research and development efforts

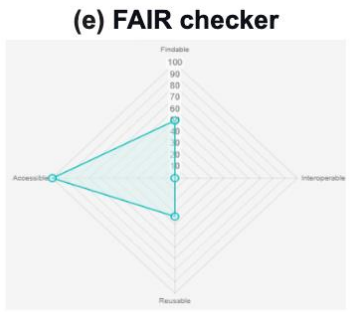


- Learning from things done in other domains
- Open source shared technology is being developed (https://hal.science/hal-04088537)
- OntoPortal is now instantiated in multiple domains and communities
- Possible out-of-the-shelf EOSC service for any community/project to deploy a Semantic Artefact Catalogue at the click of the mouse?

Background: 50 shades of FAIR!

- Generic (any type of data): SHARC, FDMM, FAIR Metrics, FAIR-Aware, FAIRshake, FAIR dat, FAIR checker
- Specific to semantic artefacts
  - H2020 FAIRsFAIR deliverable,
  - Poveda et al., (2 papers, then FOOPS!)
  - 10 simple rules paper
  - DBPedia Archivo
- Specific to semantic artefacts but pre-existing FAIR
  - MIRO guidelines (Minimum Information for Reporting an Ontology)
  - MOD (Metadata for Ontology Description and Publication Ontology)
  - 5-stars for vocabularies

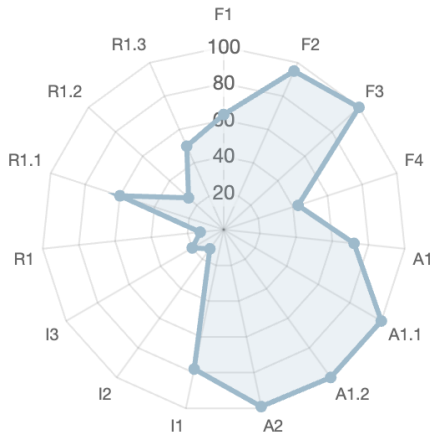




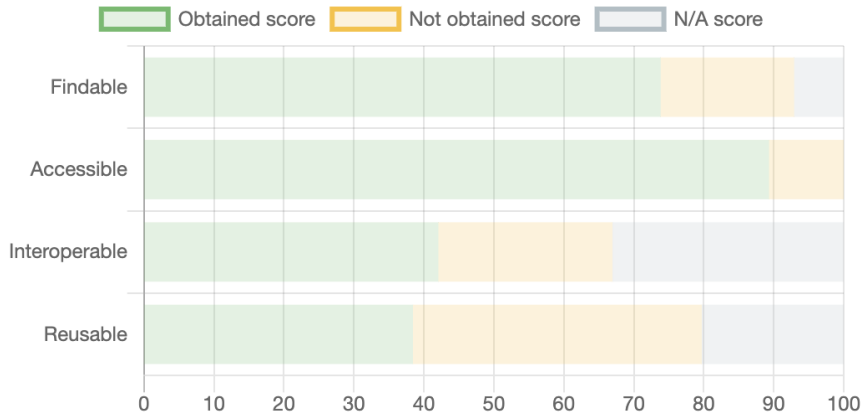
Recent work on FAIRness or alike…

- **[SHARC]** R. David et al., "Fairness literacy: The achilles' heel of applying fair principles," *Data Science Journal*, vol. 19, no. 1, pp. 1–11, Aug. 2020, doi: 10.5334/DSJ-2020-032.
- **[FDMM]** C. Bahim et al., "The FAIR data maturity model: An approach to harmonise FAIR assessments," *Data Science Journal*, vol. 19, no. 1, pp. 1–7, Oct. 2020, doi: 10.5334/DSJ-2020-041.
- **[5-star]** A. Hasnain and D. Rebholz-Schuhmann, "Assessing FAIR data principles against the 5-star open data principles," in *ESWC 2018 Satellite Events*, Jun. 2018, vol. 11155 LNCS, pp. 469–477. doi: 10.1007/978-3-319-98192-5_60.
- **[MIRO]** N. Matentzoglu, J. Malone, C. Mungall, and R. Stevens, "MIRO: guidelines for minimum information for the reporting of an ontology.," *J Biomed Semantics*, vol. 9, no. 1, p. 6, Jan. 2018, doi: 10.1186/s13326-017-0172-7.
- **[Povedal et al.]** D. Garijo and M. Poveda-Villalón, "Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web," in *Applications and Practices in Ontology Design, Extraction, and Reasoning*,. IOS Press, 2020. doi: 10.3233/SSW200034. + M. Poveda-Villalón, P. Espinoza-Arias, D. Garijo, and O. Corcho, "Coming to Terms with FAIR Ontologies," in *22nd International Conference on Knowledge Engineering and Knowledge Management, EKAW'20*, Sep. 2020, vol. 12387 LNAI, pp. 255–270. doi: 10.1007/978-3-030-61244-3_18.
- **[FAIRsFAIR]** Y. le Franc, G. Coen, J. P. Essen, L. Bonino, H. Lehväslaiho, and C. Staiger, "D2.2 FAIR Semantics: First recommendations," Mar. 2020. doi: 10.5281/zenodo.3707985.
- **[10-simple-rule] (not used at the time)** S. J. D. Coxid, A. N. Gonzalez-Beltran, B. Magagna, and M.-C. Marinescu, "Ten simple rules for making a vocabulary FAIR," PLOS Comp. Biology, June 2021, doi: 10.1371/journal.pcbi.1009041.

O'FAIRe: Ontology FAIRness Evaluator

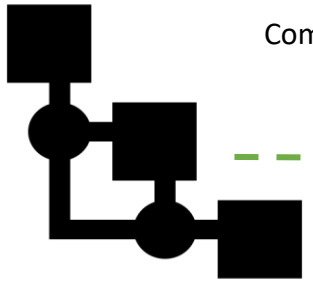


See details

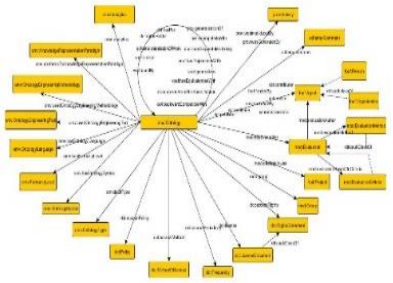


E. Amdouni, S. Bouazzouni, C. Jonquet. **O'FAIRe makes you an offer: Metadata-based Automatic FAIRness Assessment for Ontologies and Semantic Resources**. *International Journal of Metadata, Semantics and Ontologies*, 2022, 16 (1), pp.16-46. ⟨10.1504/IJMSO.2022.131133⟩.

Our objective: a **methodology** and a **tool** to automatically assess the level of FAIRness of any semantic artefacts in a catalogue



Ontologies

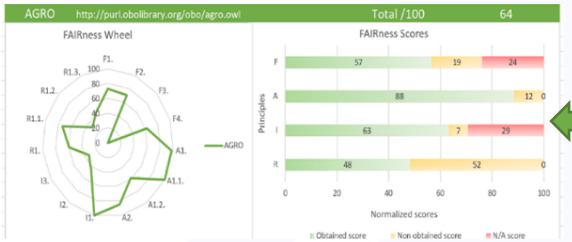


Commonn metadata model (MOD 1.4)

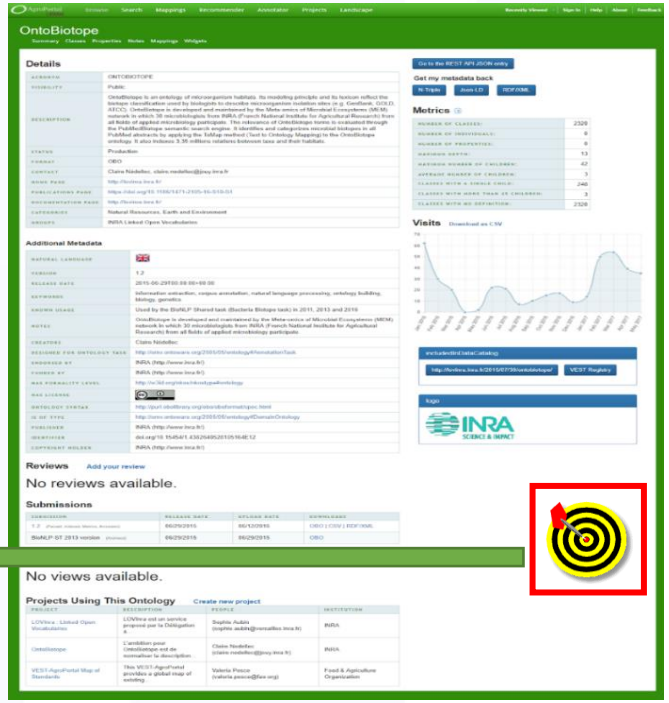
ID < … >
Auteurs < … >
Date < … >
License < … >
Type < … >
….

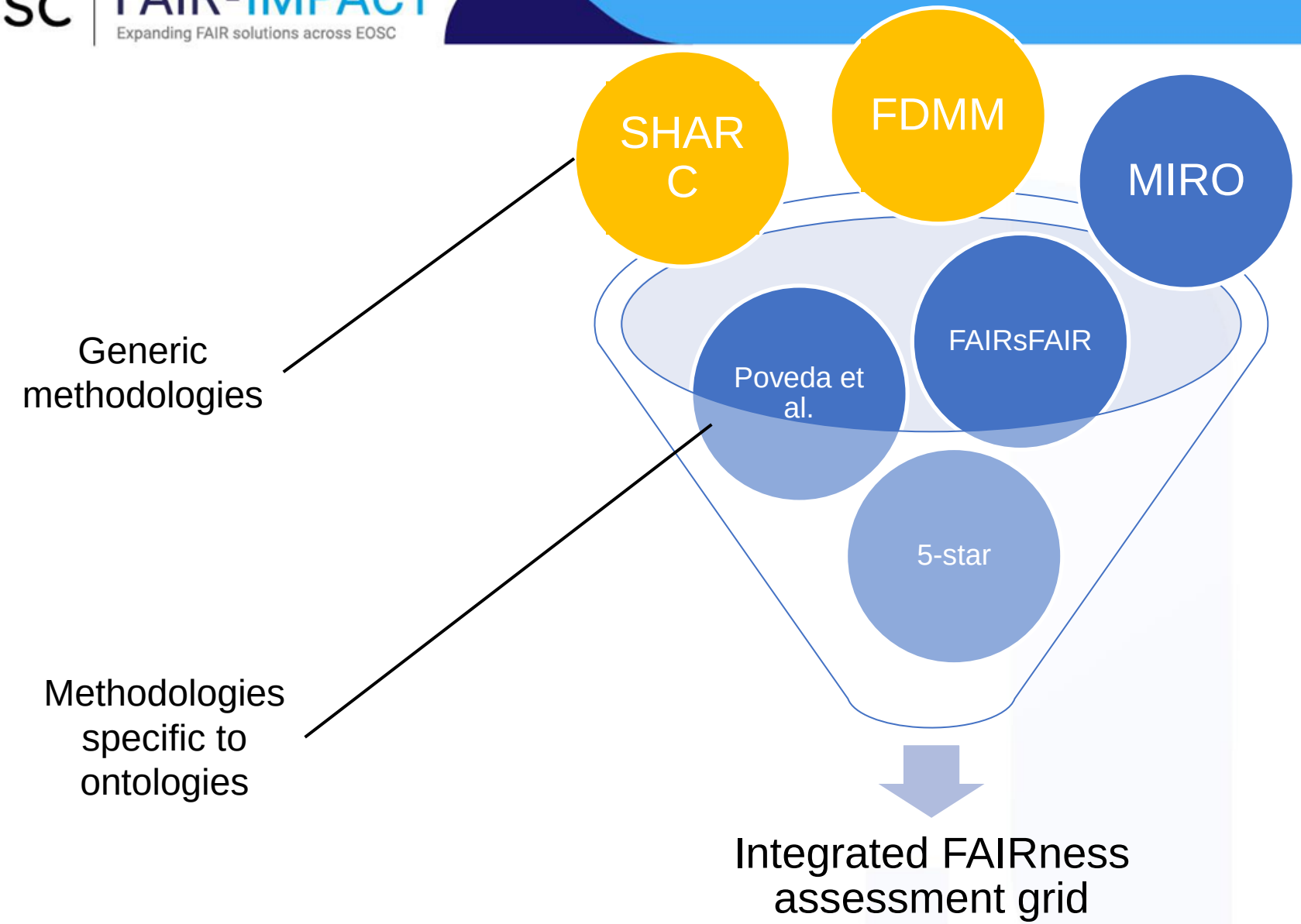
Metadata

Metadata curation



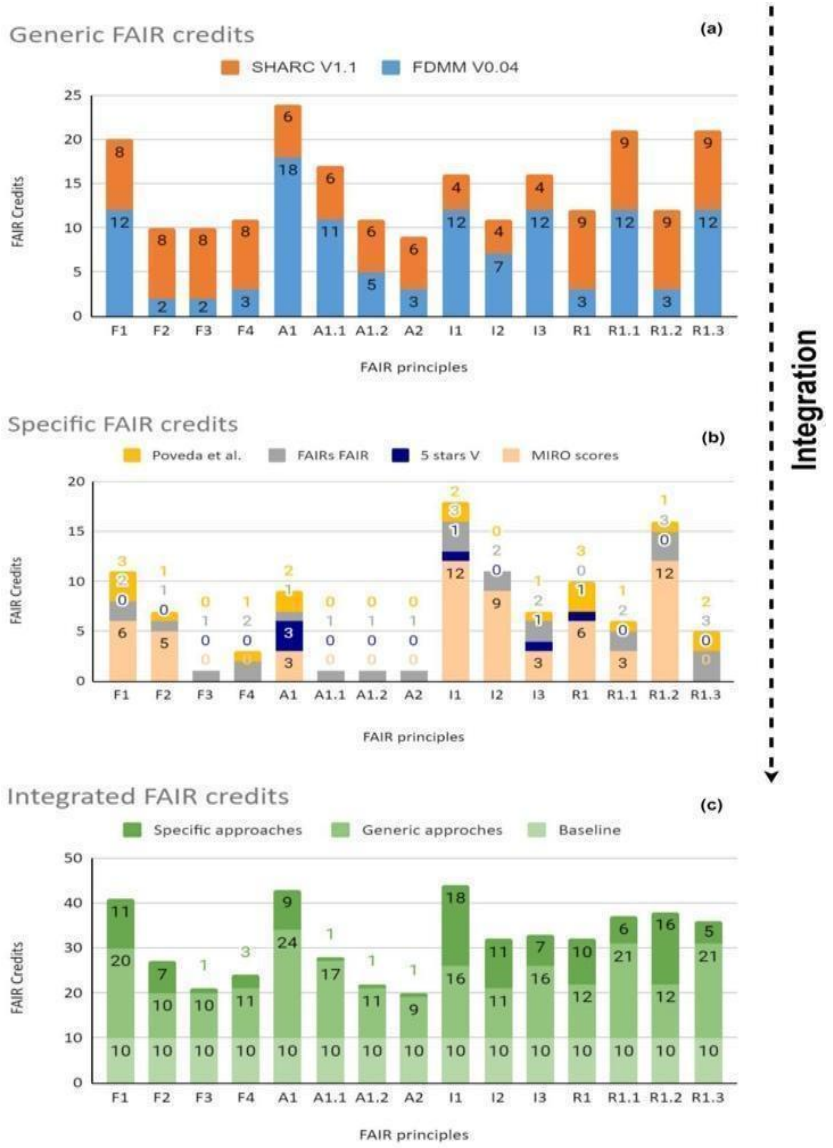
http://agroportal.lirmm.fr/




SHARC
FDMM
MIRO
FAIRsFAIR
Poveda et al.
5-star
Generic methodologies
Methodologies specific to ontologies
Integrated FAIRness assessment grid

Requirement #1: a FAIRness assessment grid

- Evaluate the importance of each principles in multiple approaches (generic & specific)
- Integrate them all in an “quantitative” grid

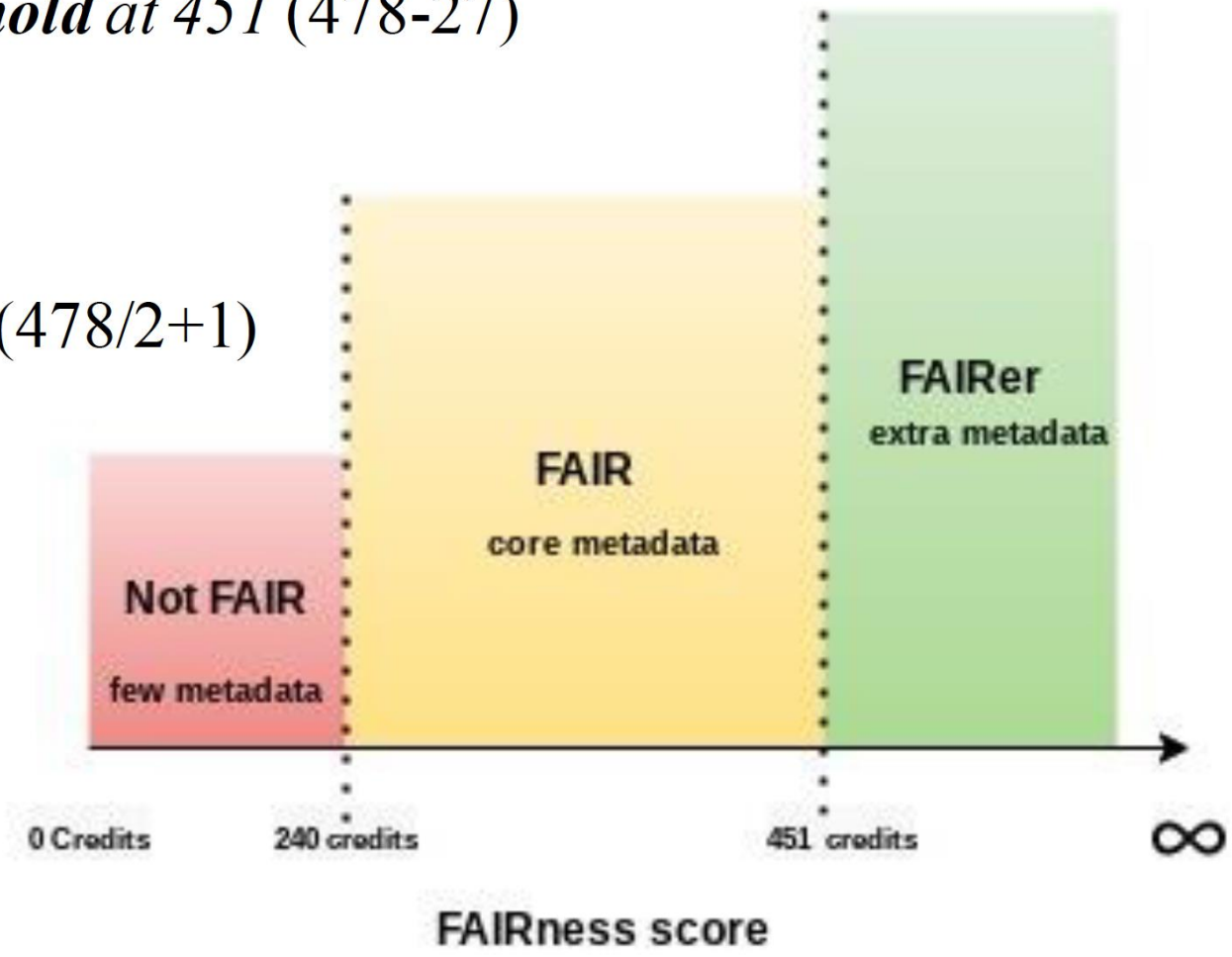


478 credits dispatched on 15 principles

FAIR or FAIRer?

***Second threshold** at 451* (478-27)
Score=94%

***First threshold** at 240* (478/2+1)
Score=50%



E. Amdouni, C. Jonquet. FAIR or FAIRer? **An integrated quantitative FAIRness assessment grid for semantic resources and ontologies**. *MTSR 2021 - 15th International Conference on Metadata and Semantics Research*, Nov 2021, Madrid, Spain. pp.67-80, ⟨10.1007/978-3-030-98876-0_6⟩.

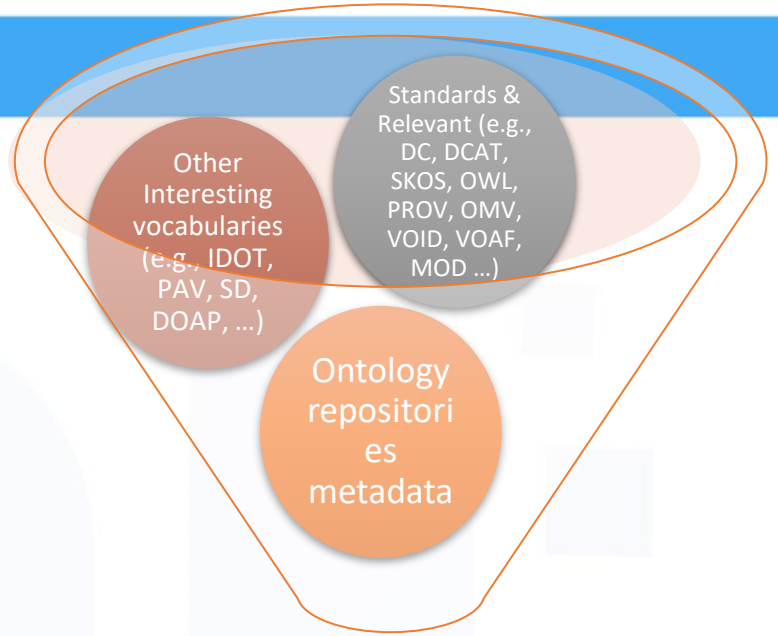
Requirement #2: a projection of the FAIR Principles for semantic resources

- FAIR Principles are very generic and need to be “projected” for different kind of research objects (cf. FAIR-IMPACT Horizon EU project)
- 61 questions
  - 45 are dependent on the semantic artefact
  - 16 are independent
- 3 examples
  - F4 Q2. Is the ontology registered in multiple open ontology 'repositories'? **10 pts**
  - A2 Q2. Are the ontology metadata of each version available? **5 pts**
  - R1.1 Q1. Is the ontology license clearly specified, with an URI that is resolvable and supports content negotiation? **15 pts**

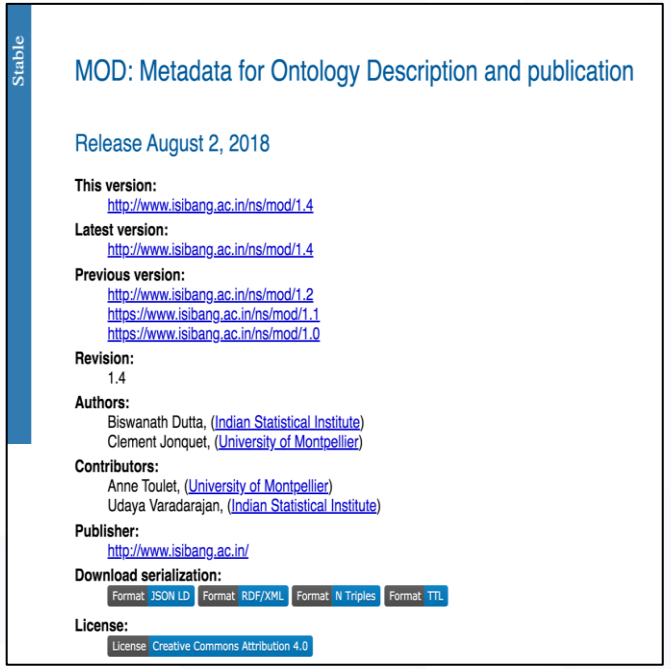
https://github.com/agroportal/fairness

$$FAIRScore\ (sr) = \sum_{j=1}^{n} FAIRSubPrincipleScore_{ij}(sr) = \sum_{k=1}^{m} QScore_{ijk}(sr)$$

Requirement #3: a unified way to describe semantic resources (metadata)



- MOD1.4 = a set of identified properties (127) one can use to describe a semantic resource
- In O'FAIRe
  - 57 MOD properties are "core" metadata properties allowing 276/478 credits
  - 70 MOD properties are "extra" metadata properties for a FAIRer level


Stable

MOD: Metadata for Ontology Description and publication

Release August 2, 2018

This version:
http://www.isibang.ac.in/ns/mod/1.4
Latest version:
http://www.isibang.ac.in/ns/mod/1.4
Previous version:
http://www.isibang.ac.in/ns/mod/1.2
https://www.isibang.ac.in/ns/mod/1.1
https://www.isibang.ac.in/ns/mod/1.0
Revision:
1.4
Authors:
Biswanath Dutta, (Indian Statistical Institute)
Clement Jonquet, (University of Montpellier)
Contributors:
Anne Toulet, (University of Montpellier)
Udaya Varadarajan, (Indian Statistical Institute)
Publisher:
http://www.isibang.ac.in/
Download serialization:
Format JSON LD Format RDF/XML Format N Triples Format TTL
License:
License Creative Commons Attribution 4.0

346 relevant properties that could be used to describe ontologies

127 used to build a new metadata model inside AgroPortal and available in MOD1.4

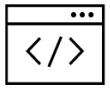
B. Dutta, A. Toulet, V. Emonet, C. Jonquet. **New Generation Metadata vocabulary for Ontology Description and Publication**. *MTSR 2017 - 11th International Conference on Metadata and Semantics Research*, Nov 2017, Tallinn, Estonia. pp.173-185, ⟨10.1007/978-3-319-70863-8_17⟩.

O'FAIRe is 80% resource metadata-based

« Findable »

- PID

owl:ontologyIRI, dct:identifier, owl:versionIRI



- Rich metadata

omv:acronym, dct:title, dct:alternative, skos:hiddenLabel, dct:description, foaf:page, omv:resourceLocator, omv:keywords …

- Metadata with PID

N/A

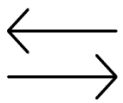
- Searchable resource

schema:includedInDataCatalog

« Accessible »

- Standardised protocol

owl:ontologyIRI, dct: identifier, sd:endpoint



- Free and open protocol

N/A

- Authentification

Schema:includedInDataCatalog

- Long term metadata access

omv:status, owl:deprecated

O'FAIRe is 80% resource metadata-based

« **Interoperable** »

- Vocabularies

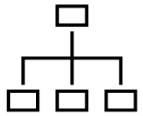
Omv:hasOntologyLanguage, omv:hasFormalityLevel, omv:hasOntologySyntax, dct:hasFormat, dct:isFormatOf

- FAIR vocabularies

F-A-I-R

owl:imports, voaf:hasEquivalenceWith, owl:priorVersion, voaf:similar, voaf:metadataVoc, dct:relation, dct:isPartOf, voaf:specializes, schema:translationOfWork,voaf:generalizes

« **Reusable** »



- Métadonnées avec attributs

mod:prefLabelProperty, mod:synonymProperty, mod:definitionProperty, mod:authorProperty, bpm:obsoleteProperty, mod:hierarchyProperty, mod:obsoleteParent, mod:maxDepth, mod:maxChildCount, mod:averageChildCount, mod:classesWithOneChild, mod:classesWithNoDefinition

- License

dct:license, dct:rightsHolder, dct:accessRights, cc:morePermissions, cc:useGuidelines

- Provenance

dct:source, prov:wasGeneratedBy, prov:wasInvalidatedBy, dct:accuralMethod, dct:accrualPeriodicity, dct:accrualPolicy, omv:versionInfo, vann:changes, dct:hasVersion, omv:usedOntologyEngineeringTool, omv:usedOntologyEngineeringMethodology, omv:conformsToKnowledgeRepresentationParadigm, omv:designedForOntologyTask, mod:competencyQuestion, dct:fundedBy

- Standards de la communauté

mod:ontologyInUse, omv:endorsedBy, mod:group, dct:accessRights

Requirement #4: a harmonized and curated environment for ontology descriptions

- AgroPortal offers a unified metadata model for every hosted semantic resources
- Metadata is curated

C. Jonquet, A. Toulet, B. Dutta, V. **Emonet. Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal**. *Journal on Data Semantics*, 2018, pp.1-31. ⟨10.1007/s13740-018-0091-5⟩.

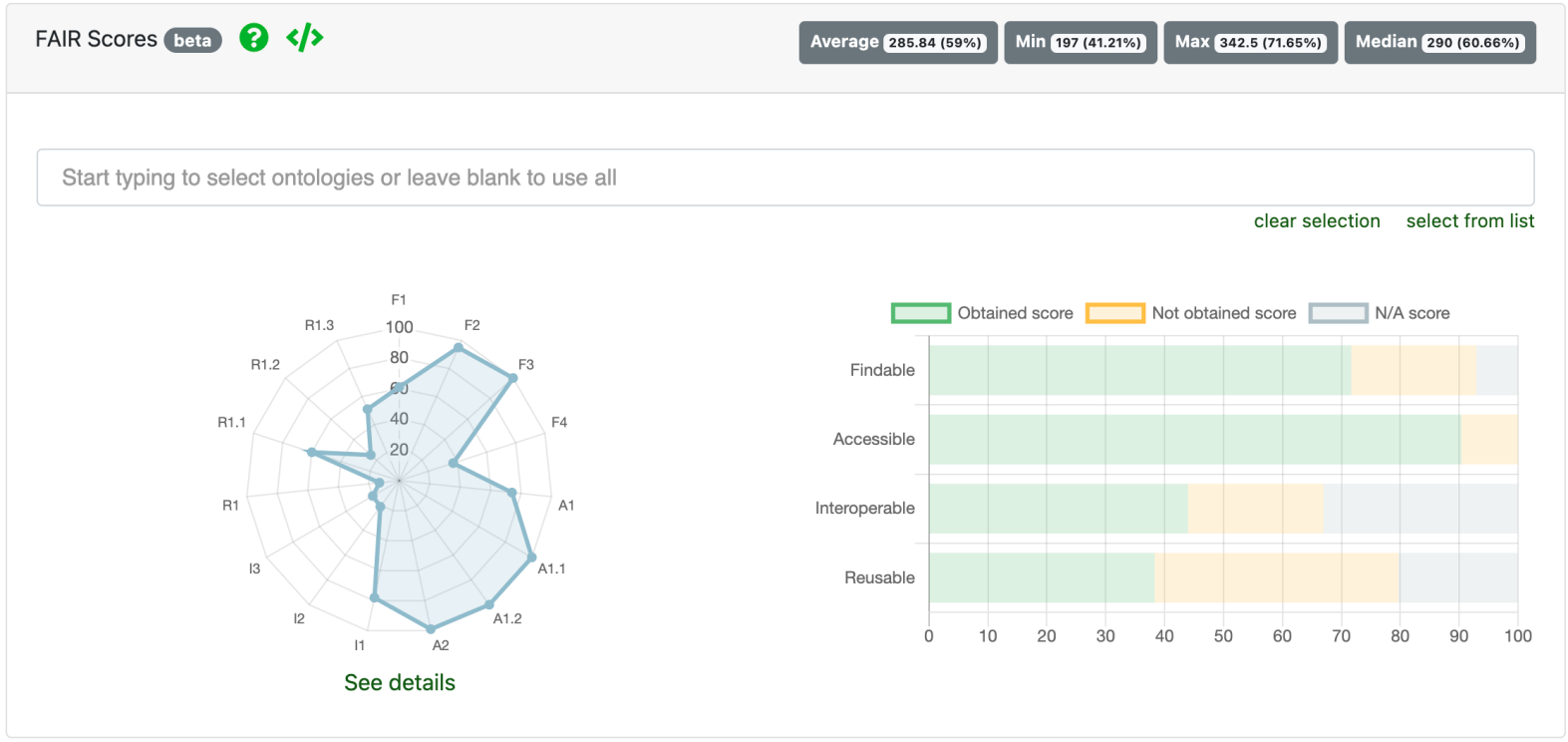
O'FAIRe: *Ontology FAIRness Evaluator*

- A methodology

  Which uses as much as possible assigned metadata values to answer a series of questions, specialized for semantic resources

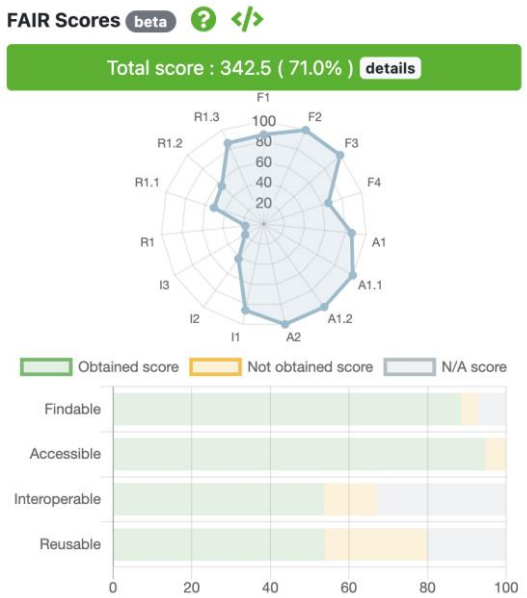
- A tool

  A web service working with any OntoPortal installations respecting the MOD1.4 metadata profile to harmonize metadata

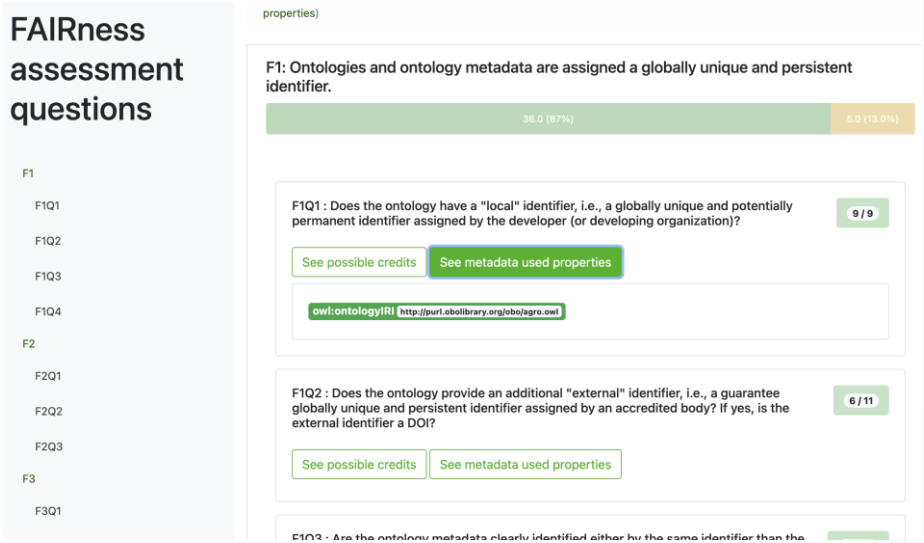


O'FAIRe in AgroPortal Demo: https://agroportal.lirmm.fr/

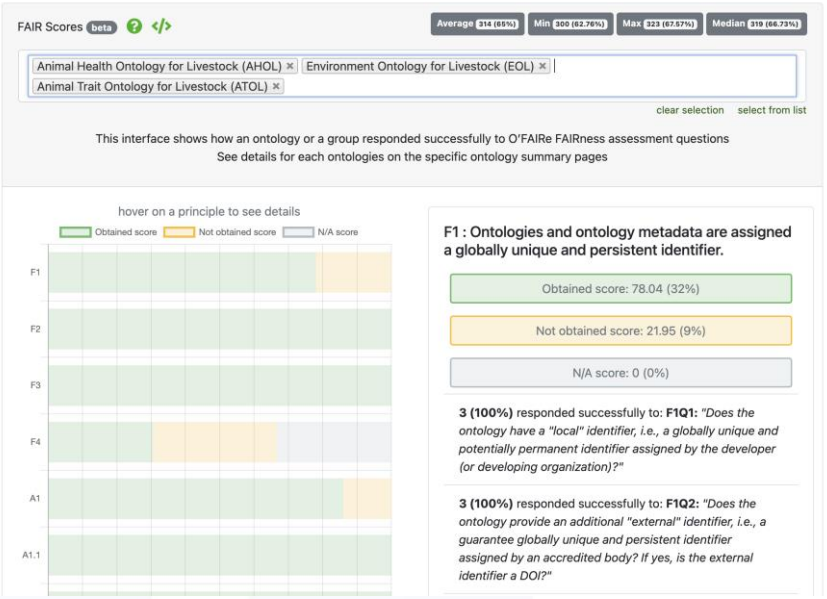
Get the FAIRness score of a given ontology



Get the explanations



Get the FAIRness score of a group of ontologies



Web service (O'FAIRe returns a JSON with the following structure)


```
{
 "ontologies": {
   "FCU": { // ontology acronym
     "Findable": { // FAIR principal
       "F1": { // Subprincipal
         "label": "Ontologies and ontology metadata are assigned a globally unique and persistent ider
         "results": {
          "F1Q1": {
                "question": "Does an ontology have a \"local\" identifier i.e., a globally unique and
                "score": 9,
                "explanation": "Present and valid ontology URI.", //Score explanation
                "properties": {//List of properties used in the test with there values
                            "owl:ontologyIRI": "http://ontology.inrae.fr/frenchcropusage"
                },
                "maxCredits": 9,
                "points": [ //Array of possible scores and explanation for this question
                       {
                         "explanation": "Ontology URI is not present.",
                         "score": 0
                       },
                       {
                       "explanation": "Present but invalid ontology URI.",
                       "score": 3
                       },
                       {
                       "explanation": "Present and valid ontology URI.",
                       "score": 9
                       }
```

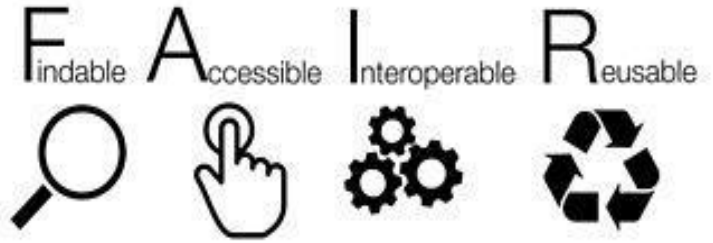

http://services.agroportal.lirmm.fr/ofaire?ontologies=AGRO

Or

http://services.agroportal.lirmm.fr/ofaire?ontologies=ATOL,EOL,AHOL&combined

Conclusion

- O'FAIRe is an approach built from other methods and contributions in the FAIR ecosystems but much more complete in terms of aspects covered
- A generic methodology with a reference evaluation grid (assuming the metadata descriptions are provided)
  - ➔ Questions can changed or be added/removed without changing the method
  - ➔ Customizable to enhance/ignore certain aspects of FAIR
- A web service working with OntoPortal ontology repositories implementing 51/61 of O'FAIRe questions
- Easier identification and selection of ontologies to use.
- FAIRer ontologies!



FAIRness assessment for semantic artefacts

Clement Jonquet (INRAE, U. Montpellier

eosc | FAIR-IMPACT

Funded by the European Union