O'FAIRe: Ontology FAIRness Evaluator in the AgroPortal semantic resource repository

Emna Amdouni, Syphax Bouazzouni, Clement Jonquet

O'FAIRe: Ontology FAIRness Evaluator in the AgroPortal semantic resource repository

Emna Amdouni[1][0000-0002-2930-5938], Syphax Bouazzouni[1][0000-0002-2069-8739] and Clement Jonquet[1,2][0000-0002-2404-1582]

[1] LIRMM, University of Montpellier & CNRS, France
[2] MISTEA, University of Montpellier, INRAE & Institut Agro, France
{emna.amdouni,syphax.bouazzouni,jonquet}@lirmm.fr

Abstract. O'FAIRe, the *Ontology FAIRness Evaluator,* is a methodology to automatically assess to which level a semantic resource or ontology respects the FAIR Principles. This paper describes the online tool implementing O'FAIRe within the AgroPortal ontology repository, through 61 questions/tests, among 80% are based on the ontology metadata description. For a specific ontology or a group of semantic resources, O'FAIRe web service outputs both global and detailed scores (normalized) against the 15 FAIR Principles. O'FAIRe results are visualized and explained with new specific user-friendly interfaces (such as the FAIRness wheel) in order to help AgroPortal users improve the FAIRness of their resources. O'FAIRe is currently implemented in three different public ontology repositories as they offer the required metadata descriptions. In the future, we will deploy the service in other OntoPortal repositories.

Keywords: FAIR Principles, FAIRness assessment, ontologies and semantic resources, ontology metadata, ontology repository.

1 Context and motivations

In 2014, the FAIR Principles established fundamental guidelines to make scientific data interoperable, persistent, and reusable for humans and machines [1]. Since then, several assessment methodologies and tools have been proposed to manually or automatically evaluate to what extent data or different research objects adhere to the FAIR Principles. For instances, FAIRdat, FAIR metrics [3], FAIRshake [2], F-UJI [3], or FAIR-checker[4]. Only one specific tool for ontologies called FOOPS! was released end of 2021 [5]. Early 2018, we argued that rich metadata descriptions and ontology repositories offer a means to facilitate the implementation of "FAIR ontologies" [6]. Later, we demonstrated the impact of harmonized and standardized metadata descriptions on the ontology identification and selection process [7]. More recently, other community efforts have also expressed the need for recommendations and guidelines on how to provide FAIR semantic resources or "artefacts" including the FAIRsFAIR H2020 project [8], or expert group guidelines [9],[10]. However, these works focus on recommendations and guidelines but do not specify a methodology for assessing the FAIRness of semantic resources (vocabularies, terminologies, thesaurus, etc.) and automating this task. FOOPS! is a good starting point for automatic FAIRness assessment, still, it has

several limits: it does not cover all the sub-principles, and does not consider and test all the related aspects of a sub-principle (e.g., "I1/I2" are evaluated with straightforward tests), and does not provide actionable guidelines to address the detected issues. It does not work with a group of ontologies. One strong difference is that FOOPS! does not rely on any ontology repository nor a standard way to describe ontologies/metadata, which is somehow both an advantage and a limitation.

From our point-of-view, clear metadata descriptions and open semantic repositories are two key elements of making semantic resources FAIR. In a previous paper, we introduced an *integrated quantitative FAIRness assessment grid* for ontologies and semantic resources [11] which dispatches 478 credits to each FAIR principle, depending on its importance when assessing the FAIRness of semantic resources. The proposed grid is based on the *Metadata for Ontology Description and Publication Ontology* [11], previous work harmonizing several metadata vocabularies into one model that has been implemented within AgroPortal [7]. With O'FAIRe, extensively presented in [REF], we go a step further and define a clear generic and customizable methodology, based on 61 questions to automatically assess the FAIRness level of ontologies, guide semantic stakeholders to make their semantic resources FAIR, and select relevant FAIR semantic resources for use. This methodology considers FAIRness assessment of ontologies should as much as possible be based on the evaluation of their metadata properties, which ones shall be ideally indexed, shared, and standardized by reference ontology repositories or libraries. As illustrated hereafter, we have implemented O'FAIRe as a web service working with any OntoPortal installation (https://ontoportal.org) [12] respecting MOD 1.4 properties and implemented specific visualizations illustrated here in the AgroPortal ontology repository [13].

2 O'FAIRe: design, implementation and demonstration

O'FAIRe is based on 61questions that describe the unambiguous tests to determine to which level a semantic resource respects a particular aspect of FAIR. The distribution of the 61 questions is as follows: Findable (13), Accessible (13), Interoperable (15), Reusable (20). Each question disposes of certain number of credits (as defined by the grid [11]) to assign to an ontology depending on how it passes the test. When assigned to an ontology, credits become points that are added and normalized into scores. The higher the number of points, the better the test is passed. For instance, for the principle R1.1 ("Ontologies and ontology metadata are released with a clear and accessible usage license."), O'FAIRe relies on 3 questions:

Q1. *Is the ontology license clearly specified, with an URI that is resolvable and supports content negotiation? 15 pts (assessed with the property* `dct:license`*).*

Q2. *Are the ontology access rights specified and permissions documented? 7 pts (assessed with the property* `dct:accessRights`*).*

Q3. *Are the ontology usage guidelines and copyright holder documented? 15 pts (assessed with the properties* `cc:morePermissions`*,* `cc:useGuidelines` *and* `dct:rightsHolder`*).*

We implemented O'FAIRe into a web service which executes tests automatically evaluating how a semantic resource stored within AgroPortal responds to the 61 questions. The tool provides a score for each sub-principles as well as a global normalized

[0-100] FAIR score. Formally speaking, we use AgroPortal's metadata record to evaluate the level of FAIRness of the corresponding semantic resource. Consequently, we do not evaluate the level of FAIRness of an ontology but the level of FAIRness of the ontology stored within AgroPortal. This distinction is important as several FAIR sub-principles are linked to the repository in which the ontology is hosted.

The questions and the web service have been implemented in a Java Servlet application, which consumes as entry the JSON ontology metadata descriptions returned by AgroPortal's web service API. The code is open-source, fully documented and available for reuse/customization on GitHub: https://github.com/agroportal/fairness. Over O'FAIREe questions: 45 are dependent of the ontology and 16 are determined simply by the fact that the ontology is stored in AgroPortal; which means the repository automatically gives 93 points to an ontology (19% of the total points). Currently, the prototype implements 50/61 questions (82%). The rest of the questions are not yet implemented because we do not have: (i) either a metadata property to store the information necessary to assess the question or (ii) implemented a mechanism to analyze the ontology content. This means that the maximum score an ontology can currently obtain in AgroPortal is 387/478 (normalized score of 81/100).

O'FAIRe prototype (v2) was released in AgroPortal v2.2 release (on 2/2/22) (http://agroportal.lirmm.fr); as well as in the SIFR BioPortal (http://bioportal.lirmm.fr), a repository of French biomedical terminology and the IndustryPortal (http://industryportal.enit.fr) developed in the context of the H2020 OntoCommons project. The three are open ontology-repositories based on the OntoPortal technology and implementing MOD 1.4. O'FAIRe web service in AgroPortal is accessible at following base URL: http://services.agroportal.lirmm.fr/ofaire. It takes as input parameter an *ontology acronym* or a list of ontology acronyms. It returns a *JSON output* which contains the FAIR scores obtained for each question aggregated by sub-principle, principle and then in total (`score`). The total score is maximized by 478 and normalized for convenience and comparison (`normalizedScore`). Every test result is justified by a short sentence (`explanation`) and when relevant the list of MOD1.4 metadata properties used (`properties`), so users may be aware of how this score was obtained.

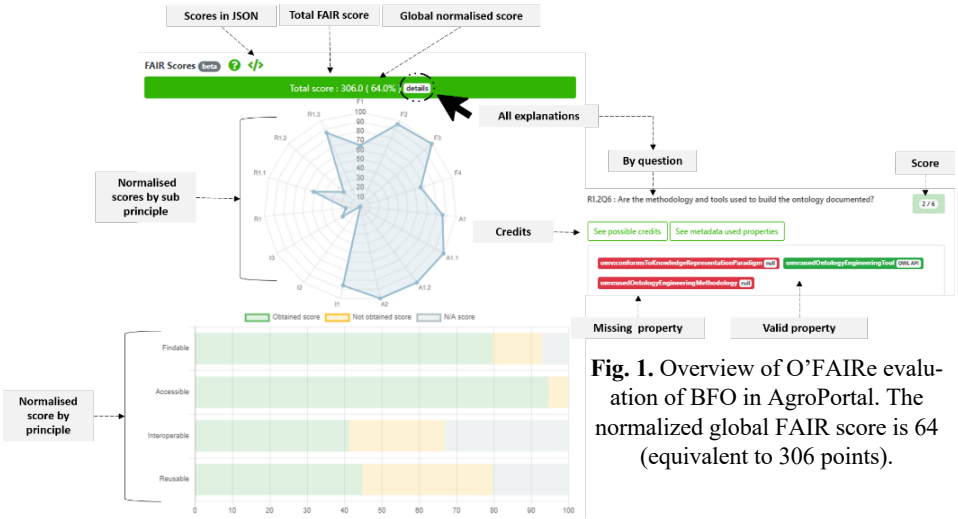

Fig. 1. Overview of O'FAIRe evaluation of BFO in AgroPortal. The normalized global FAIR score is 64 (equivalent to 306 points).

Equipped with O'FAIRe, we have revisited or developed new user interfaces within AgroPortal to display FAIR scores. For instance, it is now possible to order all the semantic resources by FAIR score on the "Browse" page, which lists all the semantic resources in AgroPortal. Fig. 1 shows an overview of the results returned for an individual evaluation of the Basic Formal Ontology (BFO) in AgroPortal: the *FAIRness wheel* shows the obtained scores over the 15 FAIR sub-principles; the *bar chart* details for each FAIR principle: the total score obtained (i.e., green part) as well as non-obtained points (yellow part) and credits that cannot yet be assigned (gray part) per limits of current implementation. Other interfaces (e.g., the *Summary page*) provides details about an ontology score, metadata properties used and explanations.

When a list of ontologies is passed as entry, the `combined` parameter computes metrics for the group of ontologies requested (average, min, max and median and returns the average scores). Fig 2. shows an illustration in AgroPortal for a group of ontologies.

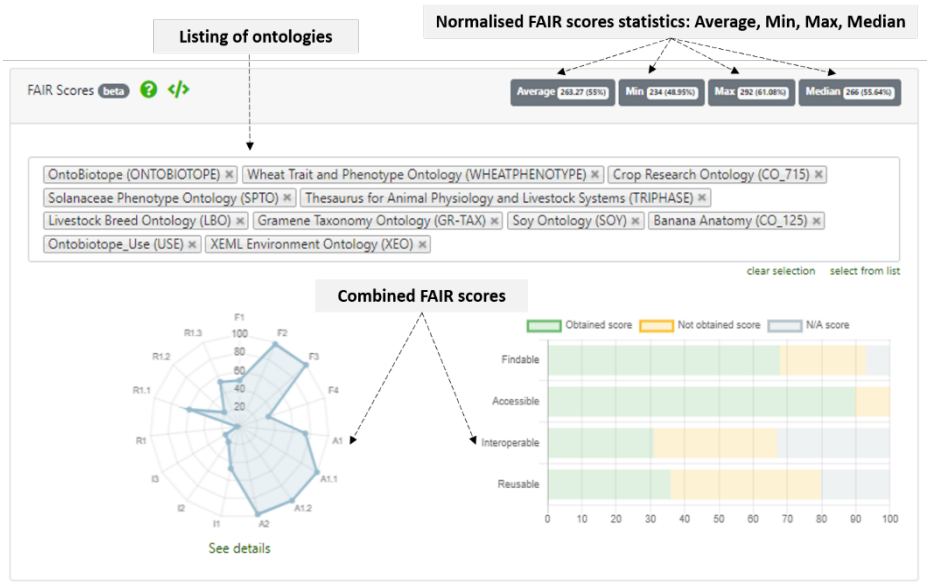

Fig. 2. O'FAIRe combined FAIRness evaluation of 11 ontologies in the OBO group in AgroPortal: average= 55, min=48, max=61, and median=55.

3 Conclusion

O'FAIRe offers both a methodology and a tool (illustrated here in AgroPortal) to enable automatic FAIRness assessment of ontologies. It differs from existing initiatives, as it is specialized for ontologies or semantic resources and it is based on metadata description harmonized in an ontology repository. O'FAIRe main goal is to offer a metric to measure the level of FAIRness and thus guide semantic stakeholders to make their semantic resources more FAIR, and select relevant FAIR semantic resources for their use. The grid on which O'FAIRe is conceived as well as its methodology (e.g., list of questions) can be customized, extended, or improved by other semantic experts in further

studies. Currently, O'FAIRe can be used in the AgroPortal an ontology repository dedicated to agronomy, the SIFR BioPortal and IndustryPortal. Collaborations within the OntoPortal Alliance will enable us to extend and maybe customize O'FAIRe for other repositories such as the BioPortal, EcoPortal or MatPortal.

Since its release, O'FAIRe was received with good interest from AgroPortal users and we have already seen some semantic resources metadata modified to 'get a better score'. In a near future, we will conduct a user survey to evaluate and improve the tool and the underlying methodology. We acknowledge that the set of questions and credit assignments are discussable and will work to reach the largest consensus in subsequent versions of O'FAIRe.

Acknowledgement

Data to Knowledge in Agronomy and Biodiversity project (D2KAB – www.d2kab.org – ANR-18-CE23-0017) and the project ANR *French participation in GO FAIR Food Systems Implementation Network* (FooSIN – https://foosin.fr – ANR19-DATA-0019).

References

[1] M. D. Wilkinson *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, vol. 3, 2016, doi: 10.1038/sdata.2016.18.

[2] D. J. B. Clarke *et al.*, "FAIRshake: Toolkit to Evaluate the FAIRness of Research Digital Resources," *Cell Systems*, vol. 9, no. 5, pp. 417–421, Nov. 2019.

[3] A. Devaraju and R. Huber, "An automated solution for measuring the progress toward FAIR research data," *Patterns*, vol. 2, no. 11, Nov. 2021, doi: 10.1016/j.patter.2021.100370.

[4] T. Rosnet, V. Lefort, M.-D. Devignes, and A. Gaignard, "FAIR-Checker, a web tool to support the findability and reusability of digital life science resources," Jul. 2021.

[5] D. Garijo, O. Corcho, and María Poveda-Villalón, "FOOPS!: An Ontology Pitfall Scanner for the FAIR principles," 20th International Semantic Web Conference, ISWC'21: Posters, Demos, and Industry Tracks, CEUR Workshop Proceedings, 2980, October 2021.

[6] C. Jonquet, "FAIR data requires FAIR ontologies, how do we do?," *F1000Res*, vol. 7, Mar. 2018, doi: 10.7490/F1000research.1115343.1.

[7] C. Jonquet, A. Toulet, B. Dutta, and V. Emonet, "Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal," *Data Semantics*, vol. 7, pp. 191–221, Aug. 2018, doi: 10.1007/s13740-018-0091-5.

[8] 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.

[9] 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.

[10] J. Malone, R. Stevens, S. Jupp, T. Hancocks, H. Parkinson, and C. Brooksbank, "Ten Simple Rules for Selecting a Bio-ontology," *PLOS Computational Biology*, vol. 12, no. 2, p. 6, 2016.

[11] B. Dutta, A. Toulet, V. Emonet, and C. Jonquet, "New Generation Metadata vocabulary for Ontology Description and Publication," in *11th Metadata and Semantics Research Conference, MTSR'17*, 2017, vol. 755. doi: 10.1007/978-3-319-70863-8_17.

[12] J. Graybeal, C. Jonquet, N. Fiore, and M. A. Musen, "Adoption of BioPortal's Ontology Registry Software: The Emerging OntoPortal Community," 2019.

[13] C. Jonquet *et al.*, "AgroPortal: A vocabulary and ontology repository for agronomy," *Computers and Electronics in Agriculture*, v144, p.126–143, 2018, doi: 10.1016/j.compag.2017.10.012.