O'FAIRe makes you an offer: Metadata-based Automatic FAIRness Assessment for Ontologies and Semantic Resources

Emna Amdouni[a], Syphax Bouazzouni[a],Clement Jonquet[a,b,*]

[a] LIRMM, University of Montpellier & CNRS, France
[b] MISTEA, University of Montpellier, INRAE & Institut Agro, France
* Corresponding author: jonquet@lirmm.fr

Abstract

FAIRness assessment evaluates the degree to which a digital object is Findable, Accessible, Interoperable, and Reusable. Here, our object of interest are semantic resources (from thesauri, terminologies, vocabularies to ontologies). Indeed, we have not yet seen a clear methodology implemented and tooled to automatically assess the level of FAIRness of semantic resources. The main objective of this work is to provide such methodology and tooling to guide semantic stakeholders for: (i) making their semantic resources FAIR through better use of standardized metadata; (ii) selecting relevant FAIR semantic resources for use. We propose a metadata-based automatic FAIRness assessment methodology for ontologies and semantic resources called *Ontology FAIRness Evaluator* (O'FAIRe). It is based on the projection of the 15 foundational FAIR principles for ontologies, and it is aligned and nourished with relevant state-of-the-art initiatives for FAIRness assessment. We propose 61 questions among 80% are based on the resource metadata descriptions and we review the standard metadata properties (taken from the MOD 1.4 ontology metadata model) that could be used to implement these metadata descriptions and improve the level of FAIRness of any semantic resource. We also demonstrate the importance of relying on ontology libraries or repositories to harmonize and harness unified metadata and thus allow FAIRness assessment. Moreover, we have implemented O'FAIRe in the AgroPortal semantic resource repository and produced a preliminary FAIRness analysis over 149 semantic resources in the agri-food/environment domain.

Keywords

automatic FAIR assessment, Semantic Web, ontologies, standardized metadata, and open repositories

1 Introduction

Making digital scientific data openly available remains a big challenge for both the scientific community and funding agencies. In this context, in 2014, a group of researchers, research institutions, and publishers –called FORCE 11– committed to making scientific data interoperable, persistent, and understandable for both humans and machines. They defined fundamental guiding principles called FAIR (for Findable, Accessible, Interoperable, and Reusable) that aim to conduct data stewards and publishers in their implementation choices [1]. Since then, the "FAIR movement" has been broadly embraced by scientists, data stakeholders, agencies, and governance bodies. Mons's book [2] defines the FAIR principles as not a standard nor a technology but are only specifications for digital objects. Their different aspects are related but need to be assessed independently.

In open science, the expression "FAIRness assessment" (or "FAIRness compliance") refers to evaluating to which degree a digital object adheres to the FAIR principles. It is complex as the FAIR principles are expressed at a very generic level and need to be expanded and projected to specific data objects to be more explicit. Despite the wide embracement of the FAIR principles, no universal FAIRness assessment criteria exist. Furthermore, some criteria are tricky to evaluate automatically and require –sometimes subjective– human expertise, for example, *meeting domain-relevant community standards* (R1.3) or *detailed provenance* (R1.2). Therefore, to enable FAIRness assessment and cover all the FAIR principles, it is preferable to distinguish what relies on human decision (e.g., which license to assign, which good practices to follow) from how to capture or represent information in a way a machine can use to evaluate FAIRness. We have adopted this approach, as explained after. The development of 'metrics' –i.e., a way to measure to which level a digital object implements FAIR principles– remains challenging. Several schemas and tools are proposed to manually or automatically assess the FAIRness of datasets, for instance: FAIR Metrics [3], [4], the FAIR Data Maturity Model (FDMM) [5], FAIRDat, FAIR-Aware [6] and more detailed later. However, we believe only the FAIR Metrics group proposes a framework with some universal metrics for automatizing the evaluation of FAIR digital objects. Most other approaches merely provide FAIR scoring based on forms or self-assessment questionnaires and recommendations on how to improve FAIR scores.

Semantic technologies are essential in making data interoperable and reusable, as required by "I" and "R" principles. In fact, standard vocabularies or ontologies are key to achieving a high level of FAIRness as stated by the "I2" sub-principle: *(Meta)data use vocabularies that follow FAIR principles*. However, like any other data, semantic resources and ontologies[1] –the backbone of semantic technologies– have themselves to be FAIR. Until recently, not much attention has been made to describing and evaluating ontologies using FAIR principles. Early 2018, we argued that rich metadata descriptions and ontology repositories offer a means to facilitate the implementation of "FAIR ontologies" [7]. Later, we demonstrated the impact of harmonized and standardized metadata descriptions on the ontology identification and selection process [8]. This work was accomplished in part in the context of the Research Data Alliance VSSIG[2] and the AgroPortal project, a vocabulary and ontology repository dedicated to agri-food [9]. 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 [10], or expert group guidelines [11],[12]. However, these works focus on recommendations and guidelines but do not specify a methodology for assessing the FAIRness of semantic resources and automating this task.

In a previous paper, we introduced an *integrated quantitative FAIRness assessment grid* for semantic resources [13]. This grid dispatches different credits to each FAIR principle, depending on its importance when assessing the FAIRness of semantic resources. In this paper, we go one step beyond and provide a clear methodology, implemented and tooled to automatically assess the level of FAIRness to 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 [14], [15]. We believe that clear metadata descriptions and open semantic repositories are two mandatory elements to make semantic resources FAIR. In previous work, we have built a new harmonized metadata model for an ontology repository and demonstrated that it enhances the identification and reuse of ontologies [8]. This paper demonstrates that ontology repositories are also crucial for FAIRness assessment.

We build an automatic FAIRness assessment methodology for semantic resources. This methodology uses as much as possible assigned metadata values to answer a series of questions, specialized for ontologies and semantic resources, for each FAIR principle.[3] We define 61 machine-actionable questions that describe unambiguously the tests to assess to which level a semantic resource respects a certain aspect of FAIR. This list of questions can be adapted or completed to refine the evaluation of one specific aspect without changing the overall methodology and assessment grid. We illustrate with some examples how the questions can be answered with a metadata property or an ontology repository feature. The methodology proposed is fully generic (i.e., does not depend on the ontologies, the domain, or any specific technology) but requires ontologies to be described with an extensive set of metadata properties that go beyond what ontology developers and most repositories usually provide [16].

We implemented this methodology in a FAIRness assessment tool –called *Ontology FAIRness Evaluator* (O'FAIRe)– which executes the tests automatically, evaluating how a semantic resource responds to the questions. O'FAIRe provides a global normalized FAIR score, i.e., a number between 0 and 100 that can be detailed for each FAIR principle (F, A, I, R) and sub-principle (F1,… R1.3). In addition, the tool provides a short explanation for scores. O'FAIRe is offered to: (i) enhance the level of FAIRness of a semantic resource; (ii) help end users select FAIR ontologies and thus respect the I2 principle. O'FAIRe's assessment methodology is aligned with relevant state-of-the-art initiatives for FAIRness assessment of semantic resources or digital objects in general, and we have conceived it in a way that can be customized, extended, or improved by other semantic experts in further studies. We considered the FDMM, SHARC, 5-stars, FAIRsFAIR, and Poveda et al. [11] initiatives (as detailed in Section 2) and blend them to determine a maximum number of credits for assessing each FAIR sub-principle for a semantic resource.

We implemented O'FAIRe as a Web service working with any OntoPortal installations respecting the *Metadata for Ontology Description and Publication Ontology* (MOD1.4) metadata profile [17]. We deployed it in AgroPortal (http://agroportal.lirmm.fr) and designed specific user interfaces to experience its performance on ontologies or groups of ontologies. Those preliminary results allowed us to test and validate the accuracy of our tool and to develop appropriate visualizations to display results graphically. This paper also presents a synthesized analysis of FAIRness over 149 semantic resources in the agri-food/environment domain available in AgroPortal as of January 2022. We mainly observed that Findability and Accessibility were better addressed than Interoperability and Reusability.

The rest of this paper is organized as follows: Section 2 introduces the FAIR principles and reviews FAIRness assessment-related work (i.e., generic and specific ones. Section 3 describes steps to design our metadata-based FAIRness assessment methodology, details our approach by presenting a projection of each FAIR principle for semantic resources, lists the questions with some illustration examples, and the metadata properties. Section 4 presents the O'FAIRe prototype implemented in

[1] In this paper, we consider the terms ontologies, terminologies, thesaurus and vocabularies as a type of knowledge organization systems [56] or semantic artefacts [10] or semantic resources [57]. Those are the subjects we are interested in making FAIR. For simplicity, we will sometimes use "ontology" as an overarching word.

[2] RDA Vocabulary and Semantic Services Interest Group.

[3] With regards to the expression used by the GO FAIR community [58], we can say our methodology is a FAIR Implementation Profile (FIP) for semantic resources and it relies on FIP (meta)data related questions.

AgroPortal and then provides a synthesized analysis for 149 semantic resources. Finally, Sections 5 and 6 discuss some lessons learned from our analysis and mention some future works before concluding the paper.

2 Background on FAIRness assessment

In this section, we briefly present the FAIR principles and influential actors in the "FAIR movement." Then we present FAIRness approaches in the literature, distinguishing two categories: (i) generic recommendations, schemes, or tools for any kind of digital object (or for dataset); (ii) approaches specific for the description and assessment of ontologies or semantic resources. We review both of them chronologically.

2.1 Terminology used

We use the following terminology, in part aligned with the recent-and-being-developed *FAIR Vocabulary*,[4] for referring to "main" FAIR related concepts:

- We use the term "**FAIR principle**" to refer to each F, A, I, R guidelines group.
- We call "**FAIR sub-principle**" the lower granularity for each FAIR principle (e.g., F1, F2, F3, and F4).
- "**FAIRness assessment**" refers to the use of questions/tests to assess to which degree a semantic resource is compliant with a sub-principle or, by generalization, a principle.
- A "**FAIRness assessment question**" assign a certain number (maximized) of "**credits"** to an ontology depending on how it passes the test. When assigned, credits become "**points**." The higher the number of points, the better the test is passed.
- "**FAIRness level**" refers to how much a FAIR principle or a sub-principle is respected. this level will be represented by a "**FAIR score**" i.e., the total number of points obtained by a semantic resource for a sub-principle, principle, or overall. A score can be normalized between 0 and 100 for a more straightforward representation and comparison.

2.2 Overview of the "FAIR movement"

Hereafter, we briefly introduce the meaning of each F-A-I-R aspect: First, data are *findable* when they are sufficiently described with metadata and are registered or indexed in a searchable registry or repository. Data, metadata, and other associated resources should have a unique and persistent identifier that makes them findable and referenceable by humans and machines. Second, data are *accessible* when users can retrieve them using a universally implementable and open protocol. But this does not mean data have to be openly accessible without restrictions. Sometimes, data can be FAIR and not open. In other words, FAIR data should be associated with metadata that specifies conditions by which the data are accessible. Third, data are *interoperable* when other stakeholders can easily and in a standardized way process them without the need for specific software. The "I" principles might be considered the more challenging to accomplish and still the most important key features for FAIR. Data and metadata should be described in a formal, accessible, shared, and broadly applicable language for knowledge representation. Plus, data must themselves reuse FAIR vocabularies or ontologies and include qualified references to other data and metadata. Early on, semantic Web and linked data technologies were identified as some best candidates to use for knowledge representation, machine-readability, and interoperability on the Web, but the FAIR principles cannot be reduced to the Semantic Web [18] and are voluntarily detached from any technology stack. Finally, data are *reusable* when they are provided with explicit license and data usage information for humans and machines. They should also be associated with rich metadata and documentation that detail their provenance (data specifications, funding projects, use cases, versions, experimental processes, etc.).

The FAIR principles are only described at a foundational level and have not been well defined at the implementation level. In Europe, for example, the implementation of FAIR principles is supported, in part, by the *European Open Science Cloud* program of the European Commission, specifically by the expert group on FAIR data [19]. The GO FAIR initiative has an objective to push for the adoption of the FAIR principles, especially via the description and adoption of *FAIR Implementation Profiles* and deployments of *FAIR Data Points*. Among other international initiatives pushing for the adoption of the FAIR principles, we can cite: Australia's *F.A.I.R. policy statement*,[5] US's NIH Data Commons program,[6] the Research Data Alliance [5], [20], multiple H2020 projects including FAIRsFAIR,[7] community-specific initiatives such as the *Food System* or other GO FAIR Implementation Network.[8]

Over the last couple of years, several working groups have formed to establish specifications and norms for FAIR. For example, the *FAIR digital object* model is a specification for implementing FAIR [21]. A FAIR digital object is composed of four elements, namely, the digital object itself, identifiers, standards, and metadata. It ensures the interoperability of digital data via the use of a persistent identifier associating each data with its contextual metadata and a standard format to facilitate its reuse. Furthermore, digital objects could be set in a *FAIR ecosystem* [19] proposing services or infrastructure to use data as for example, the FAIR Digital Framework proposed in 2020. Table 1 shows the different elements of a FAIR ecosystem for semantic resources.

[4] https://peta-pico.github.io/FAIR-nanopubs/principles/index-en.html
[5] https://www.fair-access.net.au/fair-statement
[6] https://commonfund.nih.gov/commons
[7] www.fairsfair.eu
[8] www.go-fair.org/implementation-networks/overview/food-systems

Table 1. FAIR ecosystem layers [19] and corresponding elements for semantic resources.

Applications and tools proposing content-based services for stakeholders	FAIR digital objects representing data in standard formats, with identifiers and metadata	Registries and repositories suggesting discovery, accessing, storage, and preservation services (ideally certified, trustworthy)	Networking infrastructure ensuring storage and processing of data, services, and automated workflows
Knowledge-based systems or applications relying on ontologies either directly or thru repositories, e.g., CEDAR workbench, PHIS, COPO, etc.	Semantic resource encapsulated in a file typically with a knowledge representation format (OWL, SKOS, OBO, etc.) and syntax (RDF/XML, TTL, JSON-LD, etc.). Describe a URI and metadata properties (dc: creator, cc: license, etc.)	Ontology libraries (metadata only) such as OBO Foundry, FAIRsharing, BARTOC, and repositories (metadata and content services) such as BioPortal, OLS, Ontobee, AgroPortal, EcoPortal, LOV, etc.	W3C compliant specifications for the Web from HTTP and URIs to RDF, SPARQL or SHACL

2.3 *Generic FAIRness assessment approaches*

The Research Data Alliance (RDA) **SHARing Rewards and Credit (SHARC)** Interest Group, created in 2017, proposed a FAIRness assessment grid to enable researchers and other data stakeholders to evaluate FAIR implementations and provide the appropriate means for crediting and rewarding to facilitate data sharing [20]. The SHARC grid defines a set of 45 generic criteria with importance levels (essential, recommended, or desirable) evaluated by answering one of four values (never/NA, if mandatory, sometimes, always) to a question; questions are sometimes dependent on one another as in a decision tree. In 2018, the RDA **FAIR Data Maturity Model (FDMM)** Working Group published a recommendation to normalize FAIRness assessment approaches and enable comparison of their results [5]. It describes a set of 47 generic criteria derived from the FAIR principles with priorities (essential, important, or useful). Both the SHARC grid and the FDMM recommendations assumed that some FAIR principles were more important than others. We have kept this philosophy in our methodology and kept the SHARC and FDMM outputs to influence our FAIRness assessment grid.

Some FAIRness assessment tools recently appeared, including FAIRdat, FAIR metrics, OzNome 5-star tool, FAIR self-assessment, FAIR-Aware, F-UJI, FAIR cookbook. We describe some of them hereafter and illustrate several graphical visualizations of FAIRness outputs in Figure.

Data Archiving and Networked Services (DANS) developed in 2017 **FAIRdat**;[9] it is addressed to data reviewers/curators; it creates a badge scheme per principle called "FAIR profile" that evaluates on a 5-stars scale how much a dataset is compliant to each FAIR principle. For example, an evaluation result as "F4-A3-I2-R3" denotes that the dataset is easily findable, accessible under some conditions, has a low degree of interoperability, and is, on average reusable. This tool has been influenced by the Open Data Certificate[10] and Berners-Lee's 5-star scheme.[11] Considering the feedback of Fairdat's participants in the testing phase of the first pilot version, at the Open Science Fair in 2017, the questionnaire is oversimplified, some criteria are missing, and the 5-stars rating system is subjective.[12] Therefore, further developments (implement 'R' sub-principles, reformulate questions, consider other FAIR initiatives, propose an archiving FAIR profile per dataset, etc.) are planned to make an operational tool for public users. Later, DANS also produced **SATIFYD**[13] a simple self-assessment tool based on 12 questions to sensibilize dataset providers to the FAIR principles.

The **GARDIAN FAIR metrics**[14] propose a list of indicators for calculating how much a CGIAR resource, publication, or dataset meets the FAIR principles. The proposal defines 5 levels of compliance for 'F', 'A', and 'I', and it determines 'R' with a simple average calculation, i.e., 'R'= 'F'+'A'+'I'/3. For example, 'F'–level 1 concerns data with "no PID and insufficient metadata and/or documentation," level 2 concerns data with "no PID but sufficient metadata and/or documentation," level 3 concerns data with "no PID but extensive metadata and/or documentation," level 4 concerns data with "PID and extensive metadata and/or documentation," and level 5 concerns data that "publish metadata in a CGIAR center repository."

In 2017, UPM-INIA, Oxford e-research center, and GO FAIR created the **FAIR Metrics** group and published the first "general, scalable, automatable FAIRness evaluation framework" [3], [4], enabling any scientific community to define, implement and share metrics –called Maturity Indicators (MIs)– based on the community interpretation of the FAIR principles. MI creation or edition can be performed through specific templates on GitHub; they describe the MI metadata, e.g., identifier, name, to which data does it apply, what is being measured. Each MI is associated with a compliance test (implemented with the SmartAPI) to automatically check if a digital object meets the conditions of a MI; the scoring system is binary

[9] https://www.surveymonkey.com/r/fairdat
[10] https://certificates.theodi.org/en/about/badgelevels
[11] https://5stardata.info
[12] https://slideplayer.com/slide/16324780/
[13] https://satifyd.dans.knaw.nl
[14] https://bigdata.cgiar.org/resources/gardian

(pass/fail for each test). Once a YAML description of a metric is approved by the internal FAIR Metrics community, it becomes publicly available, ready for use, and extendable for other kinds of digital objects. The FAIR Metrics initiative answers several challenges related to FAIR assessment with a framework for a variety of data services and components. It proposes a community-based collection of open MIs, tests, metadata, and recommendations on how to improve FAIRness. However, we believe the evaluation framework should be extended with further improvements (such as a longitudinal record of the assessment scores for a given resource, integrating evaluation into data management plans, etc.), and a peer-review approach to evaluate the MIs and make results more reliable. As mentioned in Section 5, in the future, we plan to adopt the FAIR Metrics framework to describe and implement our semantic resource FAIRness assessment methodology and externalize it from ontology repositories.

In 2018, the **OzNome 5-star tool** [22] was published by CSIRO. It is an online survey (containing multiple-choice questions) that generates a star chart representing the resultant degree of FAIR compliance of a dataset according to some specific metrics. In addition to the FAIR principles, the authors also considered Berners-Lee's 5-star linked open data and treated aspects that are not covered by some specific tool for producing their data rating system, the FAIR self-assessment tool. Similarly, our methodology was influenced by the Linked Open Data principles because semantic resources and ontologies are frequently implemented using semantic Web technologies.

In 2019, a **FAIRness assessment tool** for data librarians and information technology staff was proposed by ARDC as a series of questions related to each FAIR principle.[15] It offers a green bar indicator that specifies the overall level of FAIRness of datasets. This prototype tool reflects ARDC'S interpretation of the FAIR principles, but, as mentioned by its authors, part of the proposed questions has been inspired by the FAIRdat and 5-star data rating tool. Based on our knowledge, no specifications about the scoring scheme are publicly available.

In 2020, the **FAIRshake** toolkit [23] was proposed to help the community develop FAIR metrics. The toolkit enables a semi-automatic evaluation and offers the visualization of FAIRness assessment as an insignia that can be embedded within digital-resources-hosting websites.

In 2020, **FAIR-Aware** [6] was proposed by DANS, DCC, and UniHB as an output of the FAIRsFAIR H2020 project. It is an online self-assessment questionnaire composed of 10 yes/no questions (3 for F., 2 for A., 1 for I., and 4 for R). Each question is associated with detailed information and links to assist users. FAIR-Aware still needs several improvements, such as offering a synthesized score, being compliant with FAIR Metrics, providing recommendations to enhance FAIRness as a result of using the questionnaire. The FAIRsFAIR project also developed the **F-UJI tool** [24],[16] a Web service for the FAIRness assessment of any digital object based on the FAIRsFAIR Data Object Assessment Metrics. The tool implementation follows the existing Web standards such as re3data and DataCite APIs, SPDX License List, RDA metadata standards Catalog, and Linked Open Vocabularies (LOV) and PID resolution services best practices. FAIRsFAIR also formed several groups interested in the FAIRness assessment of specific digital objects such as Web services or semantic artefacts (next section).

In 2021, the **FAIR cookbook**[17] is proposed under the umbrella of the FAIRplus project. It is an online resource for the Life Sciences with exemplar recipes that aim to help researchers in learning how to FAIRify their datasets. The current recipes are browsable, and new ones could be added and improved in an open manner.

In 2021, **FAIR checker** [25] is developed by the ELIXIR-France Interoperability Working group. The Web service enables the evaluation of a digital resource, it exploits semantic Web technologies to check if metadata use standards and recognized ontologies or controlled vocabularies. The tool follows the FAIR Metrics and includes detecting missing some minimal metadata for each FAIR sub-principle.

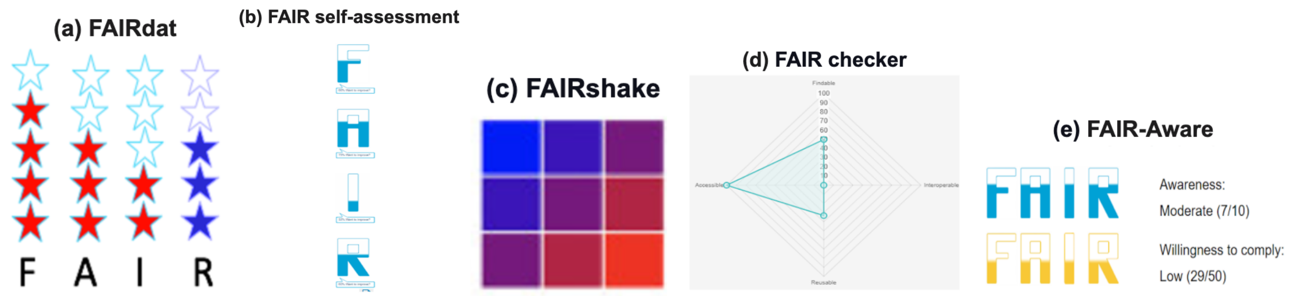


Fig. 1. Illustration of FAIRness assessment graphical visualizations from different tools: (a) *FAIRdat* uses a profile represented by 1 to 5 stars; (b) *FAIR self-assessment* uses a profile composed of blue-colored letters per principle; (c) *FAIRshake* uses a grid colored according to the level of the obtained score (from low scores in red to high scores in blue); (d) *FAIR checker* proposes a wheel visualization showing the obtained results for each FAIR principle; (e) *FAIR-Awar*e presents results with letters colored according to the level of the obtained score (in blue for awareness and yellow for willingness).

[15] https://ardc.edu.au/resources/aboutdata/fair-data
[16] https://www.f-uji.net
[17] https://faircookbook.elixir-europe.org

2.4 *Ontology specific FAIRness assessment approaches*

Before the FAIR principles, in 2011, Berners-Lee presented the foundational principles for *Linked Open Data* [26] for making data available, shareable, and interconnected on the Web. The FAIR principles have been proposed for similar reasons with a stronger emphasis on data reusability (consideration of license agreement and provenance information). The 5-stars LOD principles were specialized in 2014 for linked data vocabularies [27] as five rules to follow for creating and publishing "good" vocabularies. Under this scheme, stars denote the quality of data leading to better structure (e.g., use of W3C standards) and interoperability for reuse (e.g., metadata description, reuse of vocabularies, and alignment). The proposed 5-star rating system for vocabularies is simple; however, no implementation tool was developed for making the assessment automatic, and the principles are not largely referenced today. The degree to which the FAIR principles align and extend the 5-star open data principles was studied [11], [28] and we have incorporated this alignment in our methodology. Note that another recommendation for publishing RDF vocabularies was also produced in 2008 by the W3C Semantic Web Deployment Working Group.[18]

In 2017, the *Minimum Information for Reporting an Ontology* initiative published the MIRO guidelines for ontology developers when reporting an ontology in scientific reports [29]. The guidelines aim to improve the quality and consistency of the information content descriptions, including development methodology, provenance, and context of reuse information. These guidelines refer to 34 information items (such as "ontology name," "ontology license," "ontology URL") and specify the level of importance (must, should, optional) for each individual information item. This work was significant but was never aligned with or extended to the FAIR principles. However, the MOD 1.4 metadata model (next paragraph) provided an alignment between each MIRO guideline and the corresponding metadata properties in MOD. We have used this alignment in our methodology to influence the FAIRness assessment score with the MIRO guidelines.

Dutta et al. [17] reviewed and harmonized existing metadata vocabularies and proposed a unified *Metadata for Ontology Description and Publication Ontology* (MOD)[19] model to facilitate manual and automatic ontology descriptions, identification, and selection. MOD is not another standard nor another metadata vocabulary, but more a set of identified properties one can use to describe a semantic resource.[20] MOD 1.4 was used in AgroPortal to implement a richer, unified metadata model [8]. AgroPortal recognizes 346 properties from 15 relevant metadata vocabularies (Dublin Core, Ontology Metadata Vocabulary, VoID, FOAF, Schema.org, PROV-O, DCAT, etc.) and maps them to its unified model. This unified metadata model was the first step for enabling FAIRness assessment. For example, an ontology developer can focus on his/her responsibility of determining the license to use an ontology. At the same time, MOD and AgroPortal offer a means to encode such information in a way machines can use. Our assessment methodology relies on MOD suggested properties to implement tests as detailed in Section 3.2.

In March 2020, the FAIRsFAIR H2020 project delivered the first version of a list of 17 recommendations and 10 best practices recommendations for making semantic artefacts FAIR [10]. For each recommendation, the authors provided a detailed description, a list of related supporting technologies or technical solutions. Similarly, best practices are introduced as recommendations not directly related to a FAIR principle but contribute to the overall evaluation of a semantic resource. This proposal is currently being discussed outside of the FAIRsFAIR projects, especially in the context of the RDA VSSIG. The recommendations are also publicly available for comments on GitHub.[21] Our group is currently strongly involved in revising and commenting on the final version of the recommendations.

Later in 2020, Garijo et al. [11] considered some of the works already cited and produced "guidelines and best practices for creating accessible, understandable and reusable ontologies on the Web." In another position paper, Poveda et al. [30] complete their work with a qualitative analysis of how four ontology publication initiatives cover the foundational FAIR principles. They propose some recommendations on making ontologies FAIR and list some open issues that might be addressed by the semantic Web community in the future. These two publications are very relevant for our methodology; our work is a step further as it completes this work and proposes a concrete implementation for automatic assessment. In October 2021, Garijo et al. proposed FOOPS! a Web service for assessing an ontology regarding the FAIR principles [31]. The tool is a good starting point for automatic assessment, and it adopts, like ours, an approach based on 24 questions/tests. Still, it has several limits: it is not reusing existing work covering all the sub-principles, it does not cover all the sub-principles (e.g., missing R1.3), it does not propose actionable guidelines to address the detected issues, and it does not enable the assessment of a set of ontologies. One big difference is that FOOPS! does not depend/rely on any ontology repository nor a standard way to describe ontologies/metadata, which is somehow both an advantage and an inconvenient.

Other related work on FAIR principles for semantic resources exist but are not reused in our current methodology, include:

- A list of functional metrics and recommendations for *Linked Open Data Knowledge Organization Systems* (LOD KOS) proposed in 2020 [32] for assessing the functionality against FIT (Functional, Impactful,

[18] https://www.w3.org/TR/swbp-vocab-pub
[19] https://github.com/sifrproject/MOD-Ontology
[20] For instance, MOD does not require the use of a specific authorship property but rather encodes that `dc:creator`; `schema:author`, `foaf:maker`, or `pav:createdBy` can be used to say so.
[21] https://github.com/FAIRsFAIR-Project/FAIRSemantics/issues

Transformable) metrics and four recommendations for enhancing their FAIRness level (one recommendation per FAIR principle). This initiative is interesting, but the proposed preliminary recommendations are very limited; they do not cover all FAIR sub-principles e.g., Findable has only one recommendation that stresses the need "to enrich metadata as much as possible to enable data discovery" without considering other aspects such as identifiers and libraries/repositories.
- Late 2020, Cox et al. proposed guidelines ("10 simple rules") for making a vocabulary FAIR[22] and transform vocabularies that are not available following Web standards [33]. The authors do not explain how the proposed rules are aligned to each FAIR sub-principle and do not consider any related initiatives.
- DBPedia Archivo [34] is an ontology archive released end 2020 to help developers and consumers implement FAIRer ontologies. At this moment, Archivo contains about 1032 ontologies. The prototype[23] automatically discovers, downloads, archives, and rates new ontologies. Once an ontology is saved, Archivo determines its 4-stars FAIR rating, tracks its changes, and updates its scores. This work highlights the role that ontology libraries and repositories play in the FAIRification process. Unfortunately, this work is not inspired by existing research methodologies/tools. We think that it needs to be improved in order to make its 4-star rating system clearer for the community.

$$FAIRScore(sr) = \sum^{4}{}_{i=1} FAIRPrincipleScore_i(sr) = \sum^{n}{}_{j=1} FAIRSubPrincipleScore_{ij}(sr) = \sum^{m}{}_{k=1} QScore_{ijk}(sr)$$

where n is the number of sub-principles associated with a principle, and m is the number of questions associated with a sub-principle. $QScore(sr)$ is either the maximum number of credits or less depending how the sr passes the test. Scores are normalized using as:

$$NormalizedScore\ (sr) = (Score/\ Credits) * 100$$

where the credits are detailed in the grid proposed in [11].

Fig. 2. Formal FAIR score definition.

3 FAIRness assessment methodology

3.1 FAIR scores specification and evaluation

FAIRness assessment is about constructing a scoring function, here called $FAIRScore$, which for each semantic resource (sr) returns a numeric value representing the level of FAIRness. Formally, $FAIRScore(sr)$ is the sum of four $FAIRPrincipleScore$ obtained for each principle themselves by suming $FAIRSubPrincipleScore$ which are the sum of a series of $QScore$ obtained for each assessment question (Fig. 2).

To determine the appropriate credits for each FAIR sub-principle, we analyzed and merged relevant related approaches cited in Section 2.4, namely FDMM (v0.04), SHARC (v1.1), LOD 5-stars V, MIRO, FAIRsFAIR recommendations, and Poveda et al.'s guidelines. We consider both generic and specific approaches to provide a specialized methodology for ontologies still being influenced by more general concerns, as ontologies are a kind of digital object. The integration was not straightforward because none of the approaches used is simply and strictly aligned with the 15 sub-principles (e.g., FDMM provides 47 criteria), and two of them (i.e., MIRO and 5-stars V) were totally disconnected from the FAIR prism. The outcome of this alignment work is an integrated FAIRness assessment grid [13] of 478 credits that enables the evaluation of any semantic resource by creating a metric –and thus possible thresholds (Fig. 3, Annex 1). The grid is conceived in a way that can be customized, extended, or improved by other semantic experts in further studies. Next section, we explain how we dispatch the grid's credits on a series of 61 questions to assess each FAIR sub-principle in the O'FAIRe methodology.

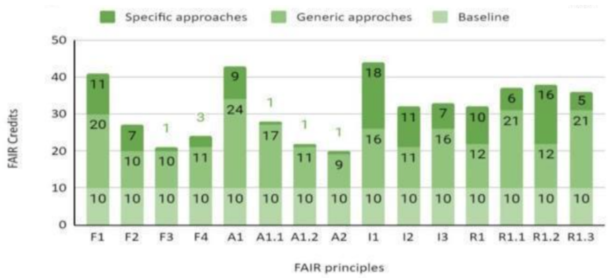


Fig. 3. Integrated quantitative FAIRness assessment grid.

3.2 FAIRness assessment questions and relevant metadata properties

In this section, we interpret each FAIR sub-principle for semantic resources and list standardized metadata properties that could be used to encode the information necessary to address the sub-principle. For each FAIR sub-principle, we propose a series of assessment questions and identify the MOD 1.4 properties to use to answer these questions. MOD allows us to determine which property can be used unambiguously; however, our methodology is independent of MOD and only requires that information is represented in ontology metadata in whatever way.

Over the 61 questions we described two kinds of metadata: "core" and "extra" metadata properties. Referring to MOD1.4, we have identified 57 "core" metadata properties allowing 276 credits over a total of 478. To

[22] https://fairvocabularies.github.io/makeVocabularyFAIR

[23] https://archivo.dbpedia.org

determine the list of extra metadata properties, we separate the metadata properties for any principles from those for F2, which has to be treated apart. Indeed, F2 ("Data are described with rich metadata") was assigned all the properties that MOD1.4 include as relevant for ontologies but have not been assigned to another sub-principle (71 "extra" metadata properties). The idea is that any ontology using some of the "extra" metadata properties in addition to the core 57 ones, will be "FAIRer".

In what follows, we rephrase the original FAIR principles with the word "ontologies" (for *data*) and "*ontology metadata*" (for metadata). This helps "projecting" the FAIR principles to ontologies. We detail what this sub-principle means or implies for ontologies and ontology metadata. And for each sub-principle, we list in boxes the questions used for FAIRness assessment and the number of credits O'FAIRe will attribute if the test is valid. For each sub-principle, we have assigned every question a certain number of credits so that the total will correspond to the number of credits identified in Fig. 3 (478 total). The distribution of the credits on each question is done as much as possible "equally & logically" but may be biased by our interpretation of how to make semantic resources FAIR; however, it can easily be amended or adjusted if questions are removed/added in the future. When possible, for each sub-principle, we also briefly mention how it links to related work. As much as possible, we provide examples of how specific ontologies or semantic resources address a sub-principle. And if a sub-principle is addressed by the ontology repository or library, we provide examples from AgroPortal.

F1. Ontologies and ontology metadata are assigned a globally unique and persistent identifier. F1 is about identifiers; although it was not mentioned in the original principle expression, several studies argued that identifiers must also be resolvable, i.e., the identifier must link back to access to the resource or the resource description. FAIRsFAIR recommendations refer to the concept of *Globally Unique, Persistent, and Resolvable Identifier (GUPRI)* [35] and proposed adopting it for both ontologies and ontology metadata. Typically, ontologies described with semantic Web languages are assigned a *Unique Resource Identifier (URI)*, and metadata are either included within the ontology file or edited in a separate file. URIs are usually globally unique, but their persistent and resolvable characteristics are not guaranteed. In some cases, ontologies can be assigned an additional *Globally Unique IDentifier (GUID)* assigned by an external organization such as a DOI. Sometimes, URIs take the form of *Persistent URLs (PURLs)* that are committed to be persistent over time. Ideally, ontology URIs should be resolvable, i.e., link back to the actual resource in a relevant format to ensure a higher degree of compliance for F1.[24] When stored in a separated file, the same rules –i.e., using GUPRIs– must apply to ontology metadata file(s) as detailed later (F2 and F3).

In addition, the ontology community sometimes tries to maintain a coherent use of acronyms to identify ontologies. For instance, in the OBO Foundry[25], a mandatory short name is assigned and used for human-friendly identification of the ontology and PURL assignment.

With the MOD ontology metadata model, F1 can be assessed by checking the value assigned to the property `owl:ontologyIRI`, used to encode the ontology's URI, and the property `dct:identifier`, used to encode another "external" identifier. In addition, the property `owl:versionIRI`, which stores a version-specific URI, can also be used to evaluate if the ontology clearly distinguishes version identifiers. Ontology acronyms (encoded in MOD with `omv:acronym`) are useful information. Still, they cannot be used for FAIRness assessment as they do not respect any of the required attributes for GUPRIs. The next box summarizes the evaluation questions regarding the different types of ontology and metadata identifiers cited above – resolvability will be evaluated in A1– and specifies the proposed distribution of credits:

F1 assessment questions (41 credits)	**Identifiers**
Q1. Does the ontology have a "local" identifier, i.e., a globally unique and potentially permanent identifier assigned by the developer (or developing organization)? **9 pts**	
Q2. Does the ontology provide an additional "external" identifier, i.e., a guarantee globally unique and persistent identifier assigned by an accredited body? If yes, is the external identifier a DOI? **11 pts**	
Q3. Are the ontology metadata clearly identified either by the same identifier as the ontology (if included in the ontology file) or with its own globally unique and persistent identifier? **12 pts**	
Q4. Does the ontology provide a version-specific URI, and is this URI resolvable? **9 pts**	
Related recommendations: (i) FAIRsFAIR: P-Rec1 and 2; (ii) Poveda el al.: Rec 1, 2, and 3.	

Examples:

- The Animal Trait Ontology for Livestock (ATOL) is identified by its URI (`owl:ontologyIRI=`http://opendata.inra.fr/ATOL) but in addition has been assigned by INRAE's official data repository a specific DOI (`dct:identifier=`10.15454/1.4690062322351956E12) for long term identification of the resource. ATOL's metadata are included in the main ontology file.
- The AGROVOC thesaurus is identified by its URI (`owl:ontologyIRI=`http://aims.fao.org/aos/agrovoc) but no other external identifier. AGROVOC's metadata are not included in the main file but live aside in their own VoID file and identified with its own URI (http://aims.fao.org/aos/agrovoc/void.ttl).
- The Cell Ontology has been assigned a short name (CL) and corresponding PURL as URI by the OBO-F (`owl:ontologyIRI=dct:identifier=`http://purl.obolibrary.org/obo/cl.owl), and no other identifier is needed as the PURL respects GUPRIs requirements. CL's metadata is partly included in the ontology file but also aside in a YAML document on GitHub as encouraged by OBO-F.

[24] The semantic Web community also says 'dereferenceable'.

[25] https://obofoundry.org

- The Environment Ontology (ENVO) provides a version specific URI for each version (e.g., `owl:versionIRI=` http://purl.obolibrary.org/obo/envo/releases/2021-05-14/envo.owl). The ontology URI (`owl:ontologyIRI=` http://purl.obolibrary.org/obo/envo.owl) always resolves to the latest version specific URI, which itself resolves to the actual ontology versioned file.

F2. Ontologies are described with rich ontology metadata. F2 criterion states the importance of describing ontology with an information model that formalizes their descriptions. FAIRsFAIR (P-Rec 2) recommends using "*common minimum metadata schema to describe semantic artefacts and their content,*" and the project is currently leading an activity to identify the relevant metadata properties for such a minimal model. FAIRsFAIR recommends the use of relevant metadata vocabularies and schema such as MOD, the *Ontology Metadata Vocabulary* (OMV) [36], and the *Linked Open Vocabularies* (LOV) [37]. With all the metadata properties being previously reviewed and aggregated within MOD, we know focusing on MOD allows us to respect this recommendation. F2 implies the definition of intrinsic (e.g., file format, creation date, status, syntax, etc.) or user-defined contextual information (e.g., title, description, language, comments, projects, relations, etc.). Intrinsic metadata can be either directly provided by the author of the ontology during the creation process or automatically generated. Several approaches for ontology metadata description were extensively reviewed by Jonquet et al. [8]. We believe that an ontology metadata model should as much as possible: (i) be based on relevant existing metadata vocabularies (e.g., DCAT, DC, OMV, PROV-O, etc.), especially W3C Recommendations and (ii) help to implement recommendations and guidelines such as the ones reviewed Section 2.4.

F2 can be assessed by evaluating the quantity and quality of the metadata properties used to describe an ontology. MOD recognizes 346 properties from 23 metadata vocabularies and groups them in an ontology metadata model of 128 properties. Because our FAIRness assessment methodology is mostly based on qualitative metadata evaluation, many metadata properties will be used to assess other specific FAIR sub-principles – such as the three properties listed above for F1 and the rest of this section. Consequently, when assessing F2, we will focus on metadata properties from the MOD model that are not used to assess another FAIR sub-principle. As some properties are clearly recognized as more important/informative than others, we will rely on the MIRO qualifications ("must," "should," and "optional") as factors for every remaining property in MOD. The complete list of 71 MOD1.4 metadata properties used to assess F2 is provided in Annex 2 with alignment with the MIRO guidelines. Some of these proposed metadata are also recommended by Poveda et al. (cf. Rec 6 about ontology metadata) such as `dct:title`, `dct:description`, `dct:abstract`, `dct:created`, `dct:subject`, `dct:modified`, `foaf:page`, `foaf:logo`, `foaf:depiction`, `owl:backwardCompatibleWith`, `vann:preferredNamespace`.

F2 assessment questions (27 credits)	Metadata
Q1. Is the ontology described with additional 'MIRO must' metadata properties? **16 pts** **Q2.** Is the ontology described with additional 'MIRO should' or 'optional' metadata properties? **4 pts** **Q3.** Is an ontology described with another metadata property with no explicit corresponding MIRO requirement? **7 pts** **Related recommendations:** (i) FAIRsFAIR: P-Rec 3; (ii) Poveda el al.: Rec 6.	

Example:

- ENVO OWL source file contains 12 metadata properties from multiple metadata vocabularies: OWL, DC, DCT, DOAP, OboInOwl, RDFS, FOAF. 5 of these properties maybe used to assess other sub-principles, however 7 "extra" properties may be used to assess F2: `dct:title`, `dct:description`, `foaf:homepage`, `doap:repository`, `doap:bugDatabase` as "MIRO must", `oboInOwl:default-namespace` (mapped by MOD to `vann:preferredNamespacePrefix`) as "MIRO optional" and `rdfs:comment` which has no mapping to MIRO.

F3. Ontology metadata clearly and explicitly include the identifier of the ontologies they describe. F3 criterion requires the association between an ontology and its metadata to be explicit by mentioning the ontology GUPRI in the metadata file and vice-versa. If the ontology metadata are included in the ontology file, no additional linking is required, and this sub-principle is automatically fulfilled. This sub-principle is especially relevant when ontology metadata are stored in another file or record; then, an explicit property needs to link an ontology to its metadata.

F3 can be assessed by differencing if the ontology metadata is included in the ontology file and, if not, by identifying the ad-hoc link to the ontology. As in F1, our methodology does not favor any approaches (i.e., Q1 is the same as Q2+Q3). However, we think a clarification is needed regarding the situations in which ontology publishers provide metadata in a separate file. To the best of our knowledge, there are no existing metadata vocabularies that define a property to relate an object to its metadata. None of the 23 metadata vocabularies reviewed when building MOD offers such a property. This issue is currently being addressed by the FAIR Digital Object framework work supervised by GO FAIR[26] . Eventually, the `fdo:hasMetadata` and `fdo:isMetadataOf` properties will be soon available.

F3 questions (21 credits)	Link ontology-metadata

[26] https://fairdigitalobjectframework.org/

Q1. Are the ontology metadata included and maintained in the ontology file? **21 pts** **Q2.** If not, are the ontology metadata described in an external file? **11 pts** **Q3.** Does that external file explicitly link to the ontology and vice-versa? **10 pts** **Related recommendations:** (i) FAIRsFAIR P-Rec 2.

Examples:

- AGROVOC keeps metadata aside in a specific VoID profile, which explicitly links back to the thesaurus with the `dct:source` metadata property.
- OBO Foundry ontologies are described in GitHub by YAML files;[27] the link between an ontology and its corresponding metadata file and vice-versa is explicit (i.e., each ontology has its own file in GitHub called ONT.md) but not semantically represented with a metadata vocabulary property. For example, the Phentotype And Trait Ontology (PATO) metadata are available in pato.md file of the ontology[28].

F4. Ontologies and ontology metadata are registered or indexed in a searchable resource. A searchable resource for ontologies and ontology metadata means an application and endpoint supporting several means of querying metadata or ontology content, e.g., term and relation lookup (via identifier or text-based). This aspect is particularly emphasized by several FAIRsFAIR's recommendations addressed to repositories, e.g., P-Rec 4: "*Semantic artefact and its content should be published in a trustworthy semantic repository*" but also P-Rec 5,6,7. FAIRsFAIR identifies the NCBO BioPortal [38], EBI-OLS [39], EcoPortal [40], AgroPortal [9], LOV [37], BODC NERC vocabulary service or more generic services such as Finto.fi, BARTOC, or Research Vocabularies Australia (see Annex 3 for URLs). Both FAIRsFAIR (i.e., P-Rec 5,6) and Poveda et al. (i.e., Rec 10 about "making an ontology findable on the Web" [11]) argue F4 is better realized if "semantic repositories" are properly indexed by Web search engines.

In the literature, "ontology library" include systems that help reuse or discover ontologies by simply listing them (e.g., DAML, Protégé, or DERI listings) or by offering structured metadata to describe them (e.g., FAIRSharing, OBO Foundry, BARTOC) but no service based on the content of the ontologies. The expression "ontology repository" [41] refers to Web applications (e.g., BioPortal, EBI-OLS, AgroPortal, Ontobee) with advanced content-based features (search, browsing, metadata management, visualization, mappings) accessible via user interfaces and application programming interfaces. Other terms are used in the literature to identify such resources, such as ontology registry, collection, or terminology service, or semantic repository [42]. In our methodology, we consider both ontology libraries and repositories but distinguish them to assign scores.

F4 can be assessed by counting in how many libraries (metadata indexing) or repositories (data indexing) an ontology is hosted. MOD property `schema:includedInDataCatalog` can be used to store in which library/repository a semantic resource is included. Annex 3 lists the repositories and libraries currently used in O'FAIRe. One community can also evaluate the importance of being present in this or that ontology library/repository; however, in O'FAIRe, to keep a tool as generic as possible, we have not implemented any community-specific evaluation aspect (except in R1.3). In addition, the quality of indexing the ontology metadata (resp. content) inside an ontology library (resp. repository), as well as the quality of indexing of such libraries/repositories by external Web search engine (cf. Q3), are information that are not specific to the ontology but directly depend on the libraries/repositories. Those aspects are not evaluated yet in O'FAIRe prototype presented Section 4.1.1.

F4 assessment questions (24 credits)	Repository
Q1. Is the ontology registered in multiple ontology 'libraries'? **6 pts** **Q2.** Is the ontology registered in multiple open ontology 'repositories'? **10 pts** **Q3.** Are the ontology 'libraries' or 'repositories' properly indexed by Web search engines? **8 pts** **Related recommendations:** (i) FAIRsFAIR P-Rec 4, 5 and 6 (ii) Poveda et al. Rec 10.	

Examples:

- The Comparative Data Analysis Ontology (CDAO, http://purl.obolibrary.org/obo/cdao.owl) is available in several open ontology repositories (e.g., BioPortal, AgroPortal, EBI-OLS) and also in multiple ontology libraries such OBO-F, FAIRSharing, and Agrisemantics Map of Data Standards.
- AgroPortal is a reference ontology repository in the agri-food domain based on the generic and open source OntoPortal technology (https://ontoportal.org) [43]. AgroPortal allows users to search for metadata and ontology content; both are stored in an RDF triplestore. The system also offers a Lucene index text-based search. AgroPortal and its content are very well indexed and retrievable directly via Web search engines which make ontologies even more findable. Plus, for an agri-food ontology, being in AgroPortal means that the ontologies will be automatically registered in FAIRsharing and Agrisemantics Map of Data Standards.

A1. Ontologies and ontology metadata are retrievable using a standardized communication protocol by their identifier. A1 criterion expresses the need for a mechanism enabling to retrieve –i.e., obtain the files– ontologies and ontology metadata by their identifier(s) (cf. F1). In the semantic Web world, the standard communication protocol is HTTP paired with URIs. An ontology file hosted on a Web server becomes accessible via HTTP, a

[27] http://www.obofoundry.org/faq/how-do-i-edit-metadata.html

[28] https://github.com/OBOFoundry/OBOFoundry.github.io/edit/master/ontology/pato.md

communication protocol standardized by W3C. Plus, if the URIs of an ontology are resolvable, objects of the ontology (including itself, its metadata, and its content) will be distinguishably retrievable by HTTP. Poveda et al.'s Rec 9 requires also URIs to be resolvable with content negotiation, i.e., resolve to different formats or destinations depending on the request.

This is typically the role of repositories to provide access to ontologies or ontology metadata via HTTP and/or Web services API. An ontology repository can be the destination for resolving URIs to Web pages if it provides a specific landing page for each ontology object. For instance, OBO Foundry PURLs resolve to OntoBee or AGROVOC URIs resolve to AIMS's SKOSMOS instance. If ontologies and ontology metadata are assigned other identifiers, they should be also resolvable, e.g., a DOI becomes resolvable once prefixed by "https://doi.org/".

A1 can be assessed by verifying if ontology and ontology metadata URIs or other identifiers, if exist, are resolvable via HTTP and, if they support content negotiation. See F1 for the questions/properties used to assess the existence of identifiers. Here, we go further and verify the resolvability of the identifiers. We can also evaluate if other standardized communication protocols are supported, e.g., a REST or SOAP Web service API or a SPARQL endpoint. The property `sd:endpoint` can be used to store the SPARQL endpoint that can be used to retrieve ontology content and metadata.

A1 questions (43 credits)	**Resolvable identifiers**
Q1. Do the ontology URI and other identifiers, if they exist, resolve to the ontology? **6 pts**	
Q2. Does the ontology URI (if metadata are included in the ontology file) or the external metadata URI resolves to the metadata record? **7 pts**	
Q3. Do the ontology URI and the external metadata URI (if the metadata are not included in the ontology file) support content negotiation? **24 pts**	
Q4. Are the ontology and its metadata accessible through another standard protocol such as SPARQL? **6 pts**	
Related recommendations: (i) FAIRsFAIR P-Rec 5 (ii) Poveda et al. Rec 7,9 (iii) 5-stars V principles 1, 2, 4.	

Examples:

- ENVO provides URIs as resolvable PURLs: e.g., http://purl.obolibrary.org/obo/ENVO_00000133 will resolve to an OntoBee landing page.
- EDAM ontology (http://edamontology.org) provides a negotiable URI making the ontology available in multiple formats: HTML, RDF/XML, and OBO.[29]
- The Agri-Food Experiment Ontology[30] provides an additional external identifier (a DOI): https://doi.org/10.15454/DPBMBW, which resolves to a data repository hosting the ontology.
- The ANAEE thesaurus[31] metadata in AgroPortal is retrievable in different formats such as JSON-LD.[32]

[29] See http://edamontology.org/EDAM.uris for details about URIs.
[30] http://agroportal.lirmm.fr/ontologies/AFEO

- Ontologies hosted in AgroPortal are all described with a value for `sd:endpoint`, which is either AgroPortal's SPARQL endpoint or a specific endpoint provided by the ontology developer. For instance, TAXREF-LD taxonomy has `sd:endpoint`=http://taxref.mnhn.fr/sparql.

A1.1. The protocol to retrieve ontologies and ontology metadata is open, free, and universally implementable. A1.1 criterion refines requirements for the communication protocol. As stated in A1, semantic Web technologies are based on HTTP/URI, which meet each of these requirements. Ontology repositories and libraries are mostly Web applications/databases which natively implement HTTP access to ontologies and ontology metadata. They can also support access via other open, free, and universally implementable communication protocols such as FTP or query endpoints such as SPARQL.

A1.1 assessment is very much linked to A1. We focus here on the use of HTTP and on the use of other open, free, and universally implementable protocols.

A1.1 assessment questions (28 credits)	**Protocol**
Q1. Is the ontology relying on HTTP/URIs for its identification and access mechanisms? **20 pts**	
Q2. Is the ontology access protocol open, free, and universally implementable? **4 pts.**	
Q3. If the ontology and metadata are accessible through another protocol, is that protocol open, free, and universally implementable? **4 pts.**	
Related recommendations: (i) FAIRsFAIR P-Rec 5.	

A1.2. The protocol to retrieve ontologies and ontology metadata supports authentication and authorization when an ontology has access restriction. A1.2 criterion continues to refine requirements for the communication protocol here to ensure secure access when needed. Indeed, in some cases, ontologies or semantic resources are not publicly available and require access or licensing restrictions. In that case, their URIs are usually not resolvable or rely on HTTPS for authentication and use a Web server that supports authorization. Note that ontology metadata can be publicly accessible or not independently of the ontology content itself. Typically, this is the role of an ontology libraries/repositories to provide open access to metadata and to serve ontologies privately when they support user/group accounts (authentication) and enable access to particular users (authorization). Here again, the use of an ontology repository facilitates the realization of A1.2. This is explicitly mentioned in FAIRsFAIR P-Rec 7 ("Repositories should offer a secure protocol and user access control.").

A1.2 can be assessed by evaluating if the resolution of identifiers (A1) supports authentication and then authorization. Additionally, the MOD model suggests using the property `dct:accessRights` to describe access restrictions. We can also assess if ontologies and ontology metadata are accessible in repositories that support

[31] http://agroportal.lirmm.fr/ontologies/ANAEETHES
[32] http://data.agroportal.lirmm.fr/ontologies/ANAEETHES/latest_submission?display=all

authentication and authorization. In O'FAIRe implementation, we have chosen the second solution.

A1.2 assessment questions (22 credits)	Protocol
Q1. Is the ontology accessible through a protocol that supports authentication and authorization? **11pts** **Q2.** Are the ontology metadata accessible through a protocol that supports authentication and authorization? **11 pts** **Related recommendations:** (i) FAIRsFAIR P-Rec 7.	

Examples:

- BioPortal and AgroPortal both support authentication and authorization: a user must create an account to upload ontologies or edit ontology metadata, and the portals support user profiles (e.g., librarian, administrator). Access to private ontologies can be granted on a case-by-case approach. Additionally, ontologies in BioPortal or AgroPortal can be either fully private or semi-public, i.e., browsable and queryable but not publicly downloadable. For instances:
  - The Food classification and description system (FOODEX2, http://data.food.gov.uk/codes/foodtype) provided as RDF is a private resource in AgroPortal, not openly accessible.
  - The Medical Dictionary for Regulatory Activities (https://bioportal.bioontology.org/ontologies/MEDDRA) is browsable and accessible via API in the NCBO BioPortal, but the ontology cannot be downloaded.

A2. Ontology metadata are accessible even when the ontology is no longer available. A2 criterion argues for a clear separation of an ontology and its metadata to ensure long-term availability of the metadata. Again, to fulfill A2, the role of ontology repositories or libraries, if permitted by the ontology authors, is to ensure access to archived ontologies and their metadata. Alternatively, it should serve only the metadata if the ontology itself is no longer available. This aspect is also identified by FAIRsFAIR P-Rec 8, which indicates that an ontology repository or library should offer human and machine-readable persistence policies to save ontology and ontology metadata changes. Added to this, we believe that an ontology repository must support versioning and allow publishing ontologies during the different production phases (i.e., beta, production, etc.) and clearly inform about the status of ontologies, especially if it becomes obsolete or deprecated.

A2 can be assessed by checking if all (or a significant number of) ontology versions are versioned and archived. We can also assess if ontologies are accessible in repositories or libraries that support metadata archiving. In addition, the MOD model offers the properties `omv:status` to indicate the different production phases (alpha, beta, production, retired) and `owl:deprecated` as a boolean to indicate if the ontology is deprecated.

A2 questions (20 credits)	Metadata long term access
Q1. Is the ontology accessible in a repository that supports versioning? **7 pts** **Q2.** Are the ontology metadata of each version available? **5 pts** **Q3.** Are the ontology metadata accessible even if no more versions of the ontology are available? **4 pts** **Q4.** Is the status of the ontology clearly informed? **4 pts** **Related recommendations:** (i) FAIRsFAIR: P-Rec 8.	

Examples:

- OBO Foundry suggest best practices for versioning (when/how creating a new file) and rely on GitHub for versioning content and metadata files.
- BioPortal and AgroPortal archives all the versions of the ontologies (and their metadata) previously hosted on the repository. Each version is automatically numbered and can be downloaded in different formats; however, only the content of the latest version is accessible via the repository services such as browsing, navigating, search, and annotation.
- AgroPortal supports access to rich ontology metadata even when an ontology is no longer available.

I1. Ontologies and ontology metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation. I1 criterion emphasizes the importance of the knowledge representation language, an aspect identified by most existing recommendations. An ontology is typically a digital resource that relies on a formal language built to be as much as possible machine-understandable. However, some ontologies or semantic resources might –for miscellaneous reasons– be described in textual (verbose) or graphical (image) form, which are not directly usable by a machine. Typically, an ontology is stored in a file using a dedicated *syntax* (RDF/XML, Turtle, JSON-LD) and *representation language* (OWL, SKOS, RDFS, OBO).[33] A semantic resource or artefact may have different *levels of formality* (e.g., ontology, terminology, thesaurus, vocabulary, etc. as, for example, listed in the NKOS types vocabularies[34] or OMV [36]) to which some representation language better correspond. Ontology metadata are usually represented using the same syntax and representation language as the ontology itself. However, this assumption is not always true when the metadata are stored aside from the ontology; thus, the language for knowledge representation must be assessed independently. Note that ontologies and ontology metadata files should be parsed without any issues using a parser adapted to their syntax and language.

I1 can be assessed by evaluating the level of formality and accessibility –in the sense easily understandable– of the representation language as well as how much it is shared/adopted by a community, and still broadly applicable (i.e., not task nor domain-specific) and recommended by relevant standardization bodies (i.e., W3C). In our methodology, we have considered semantic resources, and their metadata can be described using OWL,

[33] We voluntarily do not use the term "format" as it creates a confusion between the ontology syntax and representation language. We will use the term format when there is no distinction e.g., TXT, CSV or PDF.

[34] https://nkos.dublincore.org

OBO, RDFS, and SKOS representation language (independently of their syntax) as well as in CSV, XML, PDF, or TXT formats. Our classification of these 'description formats' is detailed in Annex 4. Typically, OWL is the most formal and broadly used representation language, but it is not the most accessible. For the assessment of I1, we can also evaluate the availability of the syntax descriptions and the level of formality, as well as the existence of distributions in other formats/syntaxes. In the MOD model, the representation language of an ontology is represented via the property `omv:hasOntologyLanguage`, its level of formality is described with the property `omv:hasFormalityLevel`, its syntax is described with the property `omv:hasOntologySyntax`. If a semantic resource distribution is available in another format or syntax, this information can be described with the properties `dct:hasFormat` and `dct:isFormatOf`.

I1 questions (44 credits)	Knowledge representation
Q1. What is the representation language used for the ontology and ontology metadata? (*) **20 pts** **Q2.** Is the representation language using a W3C Recommendation? **10 pts** **Q3.** Is the syntax of the ontology informed? **5 pts** **Q4.** Is the formality level of the ontology informed? **5 pts** **Q5.** Is the availability of other syntaxes/formats informed? **4 pts** (*) Scoring scale of each representation format (cf. Annex 4): (OWL, 18 pts), (SKOS, 15 pts), (RDFS, 12 pts), (OBO, 11 pts), (XML, 10 pts), (CSV, 9 pts), (TXT, 7 pts), (PDF, 5 pts). **Related recommendations:** (i) FAIRsFAIR P-Rec 9, 10, 11 and 12; (ii) Poveda et al. Rec 6 and 9; (iii) 5-stars V principle 2.	

Example:

- The Plant Experimental Conditions Ontology (PECO) is described in OWL (http://purl.obolibrary.org/obo/peco.owl) but is also available in OBO (http://purl.obolibrary.org/obo/peco.obo), its formality level is 'ontology' (http://w3id.org/nkos/nkostype#ontology), and its syntax is (http://www.w3.org/ns/formats/RDF_XML) RDF/XML. The AgroPortal entry for PECO rigorously describes each of these aspects using MOD and URIs for metadata values.

I2. Ontologies and ontology metadata use vocabularies that follow FAIR principles. For any kind of data or metadata, I2 criterion emphasizes the use of semantic resources (called here with the generic term 'vocabularies') that are themselves FAIR.[35] I2 is the main motivation of our work on FAIRness assessment of semantic resources. However, if we consider semantic resources are themselves data to which the FAIR principles apply, then I2 emphasizes the importance of reusing/relying on other vocabularies. In the realm of semantics, ontologies are encouraged to reuse other vocabularies in different manners (as listed in Laadhar et Jonquet 2020 [44]):

- formal imports (i.e., using the `owl:imports` construct), which result in every statement in the imported ontologies being present in the importing one;
- object reuses by simply reusing URIs from another ontology (assuming no new conflicting statements);
- alignments by creating qualified mappings to external URIs representing the same (or similar) objects.

One developer can also be influenced by another semantic resource, without making any explicit references to its objects. In all cases, the different kinds of relations between the semantic resource reusing and the one being reused shall be encoded in the metadata.

Ontology metadata reuse of vocabularies is more straightforward; it consists of (i) reusing a metadata property from an external metadata vocabulary (e.g., `dc:creator`, `cc:license`, `omv:hasNaturalLanguage`) and/or (ii) assigning values to a metadata property coming from another vocabulary (e.g., https://orcid.org/0000-0003-3934-0072, https://creativecommons.org/licenses/by/4.0, http://lexvo.org/id/iso639-3/eng). This latest aspect will be covered by I3. In both cases, when reusing external vocabularies, ontologies and ontology metadata should prefer open, formally recommended (e.g., W3C Recommendations, ISO, Dublin Core standards) or community-accepted vocabularies. By reusing a standard vocabulary, it plays the role of a common denominator to make ontologies and their metadata more interoperable.

I2 can be assessed by evaluating the proportional quantity (i.e., counting the number) and type of reuses of other vocabularies and by verifying the resolution of the imported/reused/aligned objects or influential vocabularies. We believe that it is important that each community define its preferred foundational ontologies (e.g., BFO, DOLCE), metadata vocabularies (DC, DCAT, PROV) or reference standard model (e.g., QUDT, OBOE, SSN/SOSA) as well as conceptual mappings guidelines (e.g., SKOS mapping properties, `owl:sameAs`, `oboInOwl:hasDbXref`). Plus, when looking at reuses one can also evaluate if the object reused are well-formed, e.g., do they include the minimum information for those terms such as suggested by MIREOT [45] or assess if the alignments are well represented (as discussed for example in [46]). The MOD metadata model does provide 20 properties to describe relations between ontologies including for imports (`omv:useImports`), reuses (`door:ontologyRelatedTo`), alignment (`door:isAlignedTo`), influence (`door:explanationEvolution`, `voaf:similar`, `voaf:generalizes`, `schema:translationOfWork`) and reuse of metadata vocabularies (`voaf:metadataVoc`). The metadata vocabularies currently checked by O'FAIRe are listed in Annex 5 with a coarse grain "4-level based FAIRness coefficient" manually assigned.

The FAIRness of reused vocabularies –if they are resolvable– has itself to be assessed using a methodology similar to the one presented in this article and will influence the FAIRness level of the semantic resources reusing these

[35] To avoid confusion, in this section we use 'ontologies' for the semantic resources using and 'vocabularies' for resources being used.

vocabularies. To the best of our knowledge, no metadata vocabularies, nor the MOD metadata model, provide a "FAIRness score" property that could be used to store the level of FAIRness of the reused vocabularies, which could obviously create a chicken and egg issue in the assessment of I2. In the future, we will rely on the AgroPortal O'FAIRe prototype (Section 4.1.1) when reused vocabularies are themselves hosted in AgroPortal and on a fixed user self-assigned value when not. We will be extending MOD to include a specific metadata property to store FAIRness values and I2 related aspects such as alignment curation state (referred by Q5 and Q6).

I2 questions (32 credits)	**FAIR vocabularies**
Q1. Does the ontology import other FAIR vocabularies? **5 pts** Q2. Does the ontology reuse terms from other FAIR vocabularies (URIs)? **5 pts** Q3. If yes, does it include the minimum information for those terms? **3 pts** Q4. Is the ontology aligned to other FAIR vocabularies? **5 pts** Q5. If yes, are those alignments well represented and to unambiguous entities? If yes, are those alignments curated? **7 pts** Q6. Does the ontology provide metadata information about the relation to or influence of other FAIR vocabularies? **2 pts** Q7. Does the ontology reuse standard and FAIR metadata vocabularies to describe its metadata?? **5 pts** **Related recommendations:** (i) FAIRsFAIR P-Rec 10 and 14.	

Examples:

- The Flora Phenotype Ontology (FLOPO, http://purl.obolibrary.org/obo/flopo.owl) explicitly imports the Plant Ontology using an `owl:imports`.
- The Agronomy Ontology (AGRO, http://purl.obolibrary.org/obo/agro.owl) is built by explicitly reusing terms from ENVO, CheBI and PATO. The developer uses the ROBOT software [47] to materialize correct imports when building the ontology file.
- The Vertebrate Trait Ontology (VT) is declared in AgroPortal (http://agroportal.lirmm.fr/ontologies/VT) to be explicitly aligned to the Animal Trait Ontology for Livestock (ATOL) but also related to the Livestock breed Ontology (LBO) and to the Livestock Product Trait (LPT) ontology.

I3. Ontologies or ontology metadata include qualified references to other (meta)data. I3 criterion refers to the assessment of qualified references; references to vocabularies are covered by I2, thus we will focus in I3 on references to other types of data (e.g., database or a database element) that we call cross-references or annotations. Cross-references are generally used to indicate an ontology object corresponding element in a (public/open) database, e.g., the XRef (`oboInOwl:hasDbXref`) properties for terms in the OBO world[46]. Annotations can be understood in a broader sense, informing that an ontology term annotates/describes a database or database element. For ontology metadata, a reference to other (meta)data is understood when the value given to a metadata property is taken from other semantic resources or databases, e.g., the value http://lexvo.org/id/iso639-3/eng given to the property `omv:hasNaturalLanguage` to indicate an ontology contains English labels. In both cases (ontology and metadata), we will say that a reference is "qualified" if: (i) the relationship is explicitly specified (e.g., property coming from a semantic resource) and (ii) the referenced object is unambiguously identified by a GUPRI to be explicitly reused by future users of the ontology.

I3 can be assessed by evaluating the proportional quantity and quality of references to other data and by verifying the resolution of the referenced objects. As in I2, it is up to a community to define the standard databases or resources that ontologies and semantic resources must refer to. Some ontologies may have been explicitly created to index or annotate a specific data resource. I3 is extremely hard to assess as the MOD metadata model does not currently offer any mechanism to inform about the existence of qualified references from ontologies to other data resources. One needs to look explicitly at the content of the ontology. However, we can easily evaluate if the values of certain metadata properties are unambiguous GUPRIs.

I3 questions (33 credits)	**Qualified references**
Q1. Does the ontology provide qualified cross-references to external resources/databases? **20 pts** **Q2.** If yes, are those cross-references well represented and to unambiguous entities? **6 pts** **Q3.** Does the ontology use valid URIs to encode some metadata values? **7 pts** **Related recommendations:** (i) FAIRsFAIR P-Rec 10,12,15; (ii) Poveda et al. Rec 7; (iii) 5-stars V principle 3.	

Examples:

- Many ontologies such as Gene Ontology, CHEBI or Protein Ontology include qualified cross-references to databases encoded with the `oboInOwl:hasDbXref` property. However, as discussed in [46], among one million cross-references in thirty OBO Foundry ontologies only 58% were resolvable to an entity. Plus, these entities varied from other ontology concepts (mappings) to various types of cross-references to database, database elements, or even curators. This study showed that better guidelines are indeed necessary to encode cross-references within ontologies.
- Within AgroPortal, the ANAEE Thesaurus is assigned unambiguous URIs for multiple metadata properties, e.g., license (URI provided by CC vocabulary), syntax (URI provided by W3C), subject (URIs provided by AGROVOC), formality level (URI provided by OMV vocabulary), etc.

R1. Ontologies and ontology metadata are richly described with a plurality of accurate and relevant attributes. For metadata, R1 criterion is mainly related to F2 and all the other criteria where metadata requirements are elicited, including R1's subcriteria, which focus on the context under which the ontologies were generated (e.g., provenance information) or can be reused (e.g., license information). Actually, there is no clear recommendation in

FAIRsFAIR nor in Poveda et al. about R1. Thus, we concentrate here on what it means for an ontology to be richly described: When an ontology is developed collaboratively, the metadata of ontology objects (classes, instances, properties) –which are not necessarily the exact same metadata as the ontology itself– should be richly described to trace the provenance of their definition/inclusion in the ontology. Objects metadata typically contains information about what it is, who created an object, when, in which version of the ontology it was created, and other information about its provenance. Sometimes, it is relevant to clarify globally in the ontology metadata which annotation properties are systematically used to capture a certain information in each ontology object. For instance, there are several ways of representing synonym labels (such as the OBO format synonym properties, or the SKOS label properties). Thus, it is a good practice to inform globally at the ontology metadata level which properties are used for this purpose when describing the ontology objects.

R1 can be assessed then by evaluating if objects of an ontology are well described (i.e., if the properties to describe objects are informed and significantly populated). The MOD metadata model offers several properties to describe which metadata properties are used for describing objects in the ontology: `mod:prefLabelProperty`, `mod:synonymProperty`, `mod:definitionProperty`, `mod:authorProperty`, and `mod:obsoleteProperty` can be used respectively to specify preferred names, synonyms, definition, author of a class and obsolete status. The definition of an object can be either textual (e.g., using the property declared in `mod:definitionProperty`) and/or logical using OWL constructs such as restrictions (e.g., OWL quantifier, cardinality or has value restrictions) or equivalent classes (e.g., an OWL named class). Moreover, MOD also includes three properties to describe the subclass (or broader/narrower) hierarchy of the semantic resource (`mod:hierarchyProperty`, `mod:obsoleteParent`, and `mod:maxDepth`) to be used respectively to inform the hierarchy, the root of an obsolete branch in the ontology, and the maximum depth of the hierarchy tree. In our methodology, we will use all the above-listed metadata properties to evaluate the richness of the ontology description and assess R1; of course, other properties could be used such as author and date to better capture ontology object provenance; we would need them in subsequent revision of MOD.

R1 questions (32 credits)	**Rich descriptions**

Q1. Does the ontology provide metadata information about how classes or concepts are defined? **5 pts**

Q2. Does the ontology provide metadata information about its hierarchy? **3 pts**

Q3. How much of the ontology objects are described with labels? **7 pts**

Q4. How much of the ontology objects are defined using a text description? **6 pts**

Q5. How much of the ontology objects are defined using a property restriction or an equivalent class? **6 pts**

Q6. How much of the ontology objects provide provenance information with annotation properties? **5 pts**

Related recommendations: none.

Example:

- The Wheat Ontology (CO_321) in AgroPortal (http://agroportal.lirmm.fr/ontologies/CO_321) defines almost all R1 properties. It declares using `skos:prefLabel` as prefered label property, `skos:altLabel` as synonym property, `rdfs:comment` as definition property, and `dc:creator` as object author property. CO_321 does not declare a property used to tag obsolete objects.

R1.1. Ontologies and ontology metadata are released with a clear and accessible usage license. R1.1 criterion is about licensing. Without being a requirement to be FAIR, it is obvious that making ontologies and ontology metadata openly and freely available facilitate reuse. Whatever type of license is being chosen, R1.1 requires a machine-readable declaration and representation of the license. The absence of an explicit license might prevent others from re-using the ontology, even if the ontology is originally intended to be shared and open. FAIRsFAIR authors encourage using open licenses, preferably using Creative Commons 4.0 licensing, and Poveda et al. recommend using the metadata properties `dct:license`. We note that the site RDF license (https://rdflicense.linkeddata.es) provides URIs and machine-understandable RDF descriptions for licenses. In addition, the Creative Commons vocabulary provides several properties for access rights and license descriptions.

R1.1 can be assessed by checking if the license and access rights information are provided and resolvable (especially the license). The MOD metadata model suggests the property `dct:license` for describing license information, and `dct:accessRights` for detailing access rights (who has access to what and to do what). In addition, it proposes properties to cover information on permissions and guidelines to use the ontology (`cc:morePermissions`, `cc:useGuidelines`) as well as the copyright holder (`dct:rightsHolder`). In our methodology, we assume ontology and its metadata are ruled by the same license – situation by default when the metadata are actually described in the same file– but of course, if not, two licenses need to be specified, and the assessment credits shared between ontology and ontology metadata.

R1.1 questions (37 credits)	**Licensing & rights**

Q1. Is the ontology license clearly specified, with an URI that is resolvable and supports content negotiation? **15 pts**

Q2. Are the ontology access rights specified and permissions documented? **7 pts**

Q3. Are the ontology usage guidelines and copyright holder documented? **15 pts**

Related recommendations: (i) FAIRsFAIR P-Rec 3 and 16 (ii) Poveda et al. Rec 6.

Example:

- AEFO in AgroPortal is explicitly described with a CC-BY4.0 license (https://creativecommons.org/licenses/by/4.0) and more permissions details are provided in the French Etalab Open License description.[36]

R1.2. Ontologies and ontology metadata are associated with detailed provenance. R1.2 criterion allows to describe all the ontology provenance not already "covered" by other FAIR principles and to "make the ontology readable for reuse purposes" as mentioned in FAIRsFAIR P-Rec 13. As any other data, ontology provenance might be described with cross-domain metadata vocabularies –such as Dublin Core, the Provenance Ontology (PROV-O) or the Provenance, Authoring, and Versioning (PAV) vocabulary– and as much as possible in a machine-readable form as it is clearly expressed in FAIRsFAIR P-Rec 17. The provenance is also informed by the correct versioning (and archiving of previous versions) of the ontologies through time and by informing which other semantic resources influenced the creation of an ontology –aspect covered by I2. Besides, ontology provenance must also contain information related to the methodology and tools used to build the ontology. As well as rational information on why the ontology was built (e.g., competency questions). Provenance may be understood broadly and cover most of the metadata descriptions of an ontology. Therefore, we focus here on some specific provenance information not already assessed as "rich metadata" in F2 or in another sub-principle. The use of standard metadata vocabularies to describe metadata/provenance is already evaluated in I2.

R1.2 can be assessed by evaluating the quantity of provenance information provided by the ontology. For this, the MOD metadata model provides several properties –often taken from PROV-O or DCTerms– for describing provenance information (`dct:source`, `prov:wasGeneratedBy`, `prov:wasInvalidatedBy`) and specify how/when objects are added to the ontology (`dct:accrualMethod`, `dct:accrualPeriodicity`, and `dct:accrualPolicy`). In addition, MOD regroups several metadata properties for versioning: `owl:versionIRI` (to store a version specific URI, evaluated in F1), `owl:versionInfo` (to store the version information), `dct:hasVersion` to refer to the ontology version, and `omv:hasPriorVersion` to refer to previous versions and `vann:changes` (to document the changes between versions). To document the methodology and tools used to build the ontology, the MOD metadata model offers the properties: `omv:usedOntologyEngineeringTool`, `omv:usedOntologyEngineeringMethodology`, and `omv:conformsToKnowledgeRepresentationParadigm`. To document the actors involved in the development of the ontology MOD suggest using the properties: `dct:creator`, `dct:contributor`, `pav:curatedBy`, `schema:translator`. And finally, to document the rationale about why the ontology was built one can use the properties: `omv:designedForOntologyTask`, `mod:competencyQuestion` but also `foaf:fundedBy` to inform which organization supported the work.

R1.2 assessment questions (38 credits)	Provenance
Q1. Does the ontology provide information about the actors involved in its development? **8 pts**	
Q2. Does the ontology provide information about its general provenance? **6 pts**	
Q3. Are the accrual methods and policy of the ontology documented? **6 pts**	
Q4. Is the ontology clearly versioned with version information and links to previous versions? **4 pts**	
Q5. Are the ontology latest changes documented? **2 pts**	
Q6. Are the methodology and tools used to build the ontology documented? **6 pts**	
Q7. Is the ontology rationale documented? **4 pts**	
Q8. Does the ontology inform about its funding organization? **2 pts**	
Related recommendations: (i) FAIRsFAIR P-Rec 3, 13, 15 and 17; (ii) Poveda el al. Rec 6.	

Example:

- The AGROVOC thesaurus in AgroPortal (http://agroportal.lirmm.fr/ontologies/AGROVOC) provides clear information about the accrual method (www.fao.org/agrovoc/maintenance), policy and periodicity (www.fao.org/agrovoc/how-add-term). It does also describe several other provenance information such as authors, contributors. The ontology engineering tool used is VocBench (http://vocbench.uniroma2.it).

R1.3. Ontologies and ontology metadata meet domain-relevant community standards. R1.3 criterion is one or the more ambiguous and harder to assess as it relates to domain-specific community practices and standards. Ontology design and engineering have been discussed by the scientific community for 30 years, and there are now multiple recognized practices and standard ways of representing things inside an ontology or any kind of semantic resource. With the emergence and specification of the Semantic Web, we now have a set of standard technologies and languages (often recognized as W3C Recommendations) available to build any kind of semantic resources along the semantic continuum: RDFS, SKOS, and OWL. In addition, several different communities have produced patterns or guidelines; see for instance respectively the Ontology Design Patterns [48] and the OBO Foundry principles [49]. When building an ontology, even if it is not mandatory because of the diversity of knowledge to capture, it is recommended to follow as much as possible, community or shared principles. Guidelines such as MIRO or metadata profiles such as MOD provide information on what and how to document when reporting an ontology.

Assessing if an ontology follows community standards is hard because one needs to look deeper inside an ontology and evaluate how it is built in relation to state-of-the-art guidelines and practices common in a community. Plus, this is the only FAIR principle that is "domain specific" i.e., for

[36] https://www.etalab.gouv.fr/wp-content/uploads/2018/11/open-licence.pdf

which a generic assessment measure cannot be implemented. In the O'FAIRe methodology, we decided to simply recognize and acknowledge the use of a standard representation language or syntax (points evaluated in I1); the recognition, use or endorsement of an ontology by a project or an organization; and the inclusion of an ontology in certain curated/guided ontology frameworks or groups such as for instances:

- The OBO Foundry (OBO-F) initiative [49] for biological/biomedical ontologies,
- The Planteome ontologies [50] for reference ontologies for plant,
- The FAO produced vocabularies for agri-food,[37]
- INRAE curated vocabularies / semantic resources in agriculture and environment,[38]
- The Crop Ontology project for plant specific trait ontologies [51],
- The ontologies endorsed by the Wheat Data Initiative [52],
- The WHO produced vocabularies for health,
- The Industry Foundry Ontologies project.

In R1.3, we will be evaluating if an ontology is included in a group of ontologies respecting certain community practices. By doing so, we delegate to a group/organization/framework the evaluation process and only rely on their decision. In an open science vision, we may also be inclined, to recognize a higher level of FAIRness for ontologies that are fully open and accessible (even if a restricted private ontology can be FAIR too). For all these, the MOD metadata model suggests the properties `mod:ontologyInUse`, `omv:endorsedBy` for describing the information about the projects or organization endorsing an ontology, `mod:group` for describing the inclusion of an ontology to a recognized group of ontologies, and the property `dct:accessRights` informs about the free and open availability of an ontology.

R1.3 questions (36 credits)	**Community standards**
Q1. Does the ontology provide information about projects using or organizations endorsing? **10 pts** **Q2.** Is the ontology included in a specific community set or group? (*) **20 pts** **Q3.** Is the ontology openly and freely available? **6 pts** (*) This question is community specific. **Related recommendations:** (i) FAIRsFAIR P-Rec 3, 12, 13; (ii) Poveda et al. Rec 6 and 10.	

Q2 is community-specific and requires expert knowledge to decide whether an ontology belongs to a group and how much a group sets/corresponds to the 'community standards.' In the case of AgroPortal, groups associate ontologies from the same project or organization (http://data.agroportal.lirmm.fr/groups). Currently, AgroPortal considers 8 groups relevant in the agri-food domain, e.g., OBO-F, Crop Ontology project, INRAE curated, and the Wheat Data Initiative ontologies. In this context, we will assign Q2 points as follows: Is the ontology included in the OBO library? 15 pts. Is the part of the main OBO Foundry ontologies? 20 pts Is the ontology included in the WHEAT, CROP or INRAE group? 10 pts

Example:

- The Plant Ontology (PO, http://purl.obolibrary.org/obo/po.owl) is considered as a reference ontology in the community; it is included in the OBO-F group; the WHEAT and CROP group.

4 Results

4.1.1 O'FAIRe prototype in AgroPortal

In this section, we describe and demonstrate the O'FAIRe prototype (v2.0) implemented with and within the AgroPortal ontology repository. We implemented the ontology FAIRness assessment methodology previously presented into a Web service which executes tests automatically evaluating how a semantic resource stored within AgroPortal responds to the questions. 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. Furthermore, the link between an ontology and its metadata is explicitly represented in AgroPortal (F3.Q2 and Q3 are verified).

The FAIRness assessment questions have been as much as possible implemented in a Java Servlet application. The application is open-source and available for reuse/customization on GitHub.[39]. 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 (e.g., F3.Q1., I2.Q3, I2.Q5, I3.Q1, I3.Q2, R1.Q3, R1.Q4, R1.Q5, R1.Q6, and R1.2.Q5) or (ii) implemented a mechanism to deal with and analyze the ontology content (e.g., F4.Q3, I2.Q2, I2.Q5, I3.Q1). This means that the maximum score an ontology can currently obtain in AgroPortal is 387/478 (normalized score of 81/100).

For the 45 questions[40] depending on the ontology, O'FAIRe consumes as entry the JSON-LD ontology descriptions as returned by AgroPortal's Web service API. For instance, the following call returns the description for the latest version of the Agronomy Ontology (AGRO):[41]

[37] http://datalab.review.fao.org/datalab/caliper
[38] https://vocabulaires-ouverts.inrae.fr
[39] https://github.com/agroportal/fairness
[40] Except A1.Q3 that is related to both metadata and repository.
[41] AgroPortal's internal metadata model implements MOD1.4 but does not always use MOD suggested properties to stay backward compatible with BioPortal. The correspondence between MOD1.4 properties and O'FAIRe used properties in AgroPortal is available.

http://data.agroportal.lirmm.fr/ontologies/AGRO/latest_submission?display=all

O'FAIRe can assess any semantic resource stored in AgroPortal independently or in batch. When several semantic resources are given as input, O'FAIRe returns the average, min, max and median scores of all the requested resources. The Web service takes as input an ontology acronym (local identifier in AgroPortal) or a list of ontology acronyms and returns a JSON output which contains the FAIR scores obtained for each question aggregated by sub-principles and principles as well as the global score for a resource (i.e., `totalScore` and `normalizedScore`). Every score obtained is justified by a short sentence (i.e., `explanation`), and the list of metadata properties used for the evaluation is given (i.e., `properties`) so that the user may be aware of how this score was obtained. An example of a Web service call for AGRO can be accessed hereafter. The score obtained as of the time of writing is 342 (normalized to 71):

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

The `combined` parameter allows us to obtain the average score of the group of ontologies given as input. For example, the following call returns the combined FAIR scores for the three related ontologies Animal Trait Ontology for Livestock (ATOL), Environment Ontology for Livestock (EOL) and Animal Health Ontology for Livestock (AHOL):

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

The complete documentation of the prototype is available on GitHub.

The Web service was developed to be compliant with any OntoPortal repository assuming they support the MOD 1.4 metadata model. The `url` parameter can be optionally used to specify in which repository (other than AgroPortal) an ontology may be found. For instance, the following call will evaluate the level of FAIRness of AGRO stored in the NCBO BioPortal:[42]

http://services.agroportal.lirmm.fr/ofaire?url=https://data.bioontology.org&ontologies=AGRO

However, because BioPortal does not implement the equivalent of the MOD model to describe ontology metadata, many properties required to assess FAIRness are not available. Therefore, the score obtained is significantly lower: 182. We are currently discussing within the OntoPortal Alliance [43] the extension of the metadata model for all the ontology repositories based on the OntoPortal technology.

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. As illustrated in Fig. 4, a new graphic component allows one to visualize the FAIR scores obtained (here for a group of three ontologies): (left) the *FAIRness wheel* shows the obtained scores over the 15 FAIR sub-principles; (right) 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.

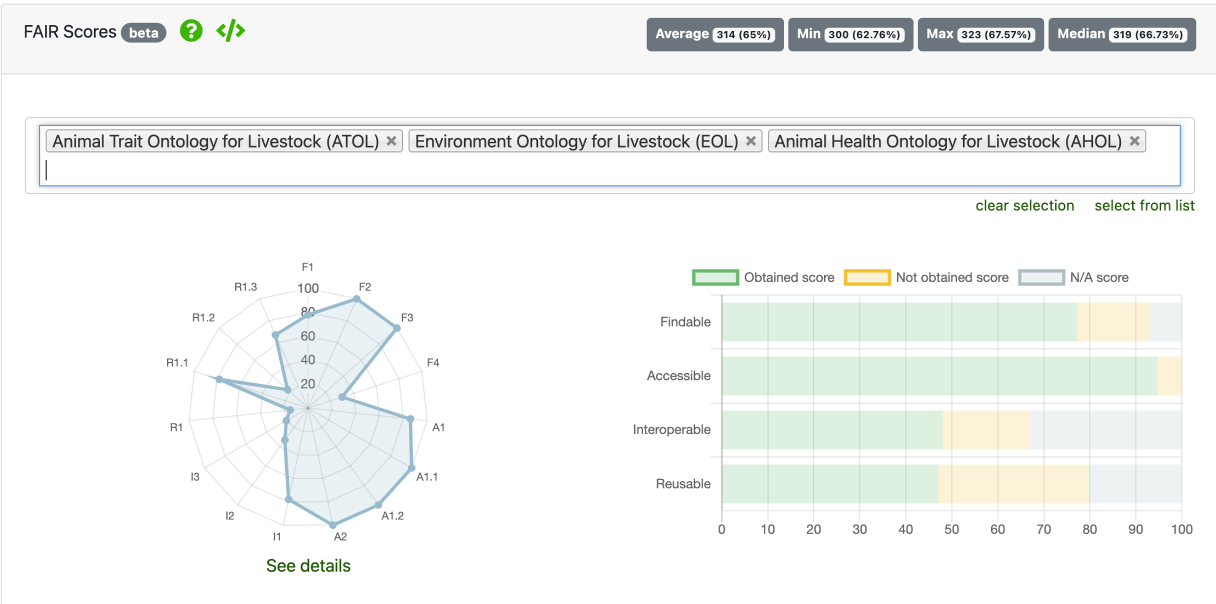


Fig. 4. O'FAIRe results in AgroPortal for the ATOL/EOL/AHOL. Presented as a FAIRness wheel (left) and bar charts (right)

[42] An API Key is required to query the NCBO BioPortal. Therefore, the REST Web service call must be completed by a parameter &apikey="…" as provided by BioPortal. The constraint exists also on AgroPortal, but in our case the system uses a default API Key to facilitate access. DURING THE REVIEWING PROCESS, ONE CAN USE THE FOLLOWING API KEY FOR NCBO BIOPORTAL: 4a482fad-9126-4696-a35a-7d42c05b82f0

4.1.2 Synthesized FAIRness analysis for semantic resources in AgroPortal

Thanks to the O'FAIRe prototype, we were able to evaluate the level of FAIRness of semantic resources stored in AgroPortal. The following discussed FAIR scores are strongly linked to the richness and quality of the metadata descriptions in AgroPortal. It is important to recall that within AgroPortal, the metadata are either entered by the ontology developers or by the AgroPortal team, which does some metadata curation/completion. Therefore, the results presented are the scores currently obtained by ontologies in AgroPortal. We expect scores will change in the future with more curation, more tests to cover all the proposed questions, and more awareness of the FAIR score produced by O'FAIRe. Our study is divided into two parts: first, we provide some statistical analysis and explore which FAIR principles are mostly 'respected' and which ones are more problematic. Second, we analyze community-specific outcomes for several groups of ontologies in AgroPortal.

We reviewed 136 public and 13 private ontologies / semantic resources hosted in AgroPortal in January 2022. The corpus includes 113 OWL ontologies, 21 SKOS thesauri, 13 OBO ontologies, and 2 terminologies from UMLS. Some of these resources are included into groups cited previously: 37 in the CROP group; 31 in the INRAE group; 24 in the OBO-F group; and 20 in the WHEAT group.

Fig. 5 presents normalized FAIR scores for AgroPortal ontologies with a minimum of 25 (this score is obtained by a private ontology that is still under revision by its ontology developers) and a maximum of 70. The average *FAIRscore* is 59. Depending on a set threshold (as discussed in [11]), we can say that almost all resources are FAIR if *FAIRScore* needs to be above 40 but only a few are if *FAIRScore* needs to be above 65. This analysis clearly shows:

- FAIR scores are relatively high (median and average around 60) which confirms that ontologies and semantic resources are digital objects that are relatively FAIR "by design" mostly because they rely on standards and technologies that obviously facilitate FAIRness. The ontology repository has also a significant role in these high scores, which confirms that depositing any kind of digital objects in data repositories make it more FAIR.
- FAIR scores distribution is relatively homogeneous with 80% of the ontologies between 50 and 65. This is mainly because the semantic resources are curated and hosted in an ontology repository and are not published in the wild, so they follow specific metadata editing and curation policies. However, we notice that the 149 semantic resources still have to be improved to fully comply with the 15 FAIR principles. Currently, the highest score is 70 on a maximum score obtainable in AgroPortal of 81. FAIR scores will unavoidably augment after implementing the missing questions in O'FAIRe; however, changes will also be required on the ontology side to reach a normalized score of 100.

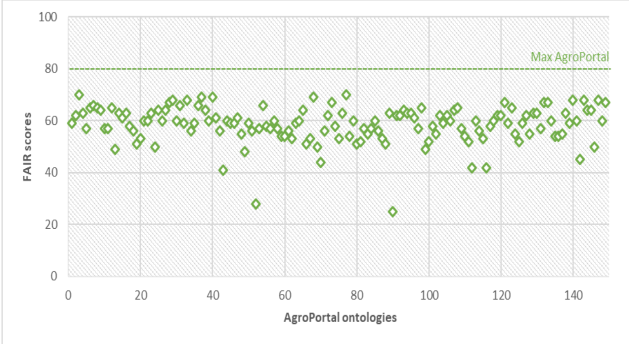


Fig. 5. The distribution of normalized *FAIRscore*s over 149 ontologies in AgroPortal (average=59, median=60, max=71, min=25).

Fig. 6 and the following paragraphs detail and discuss the average *FAIRPrincipleScores* calculated by the current version of O'FAIRe.

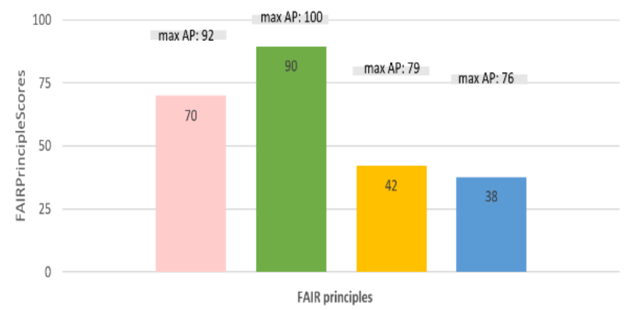


Fig. 6. Average normalized *FAIRPrincipleScores* over 149 ontologies in AgroPortal. Maximum scores obtainable within AgroPortal are mentioned above each bar.

We observe an average normalized score of 70 for F, detailed for each sub-principle in Fig. 7. When assessing F1, we observed:

- Good use of URIs as local/primary identifiers, indeed 96 ontologies (65%) provide a valid URI (Q1).
- 65 ontologies (44%) provide an external/secondary identifier which is a DOI in 10 cases (16%). Only 7 external identifiers (10%) are not resolvable, and all DOIs are resolvable (Q2).
- The repository makes explicit the relationship between the ontology and its metadata (Q3); this aspect does not discriminate ontologies if taken from AgroPortal. This is also the case in assessing F3.
- Only 21 ontologies (14%) provide a version specific URI for which 15 are resolvable (Q4).

When assessing F2, we observed almost all ontologies are described with some "rich" metadata properties: 75% of ontologies are described with at least 8 "MIRO must" metadata properties which are required to obtain Q1 points; 31% of ontologies are described with at least 6 "MIRO should and optional" metadata properties which are required to obtain Q2 points, and 98% of ontologies are described with at least 7 metadata properties with no MIRO correspondence which are required to obtain Q3 points.

When assessing F3, we have assigned 0 points to Q1, but 11 points to Q2 and 10 points to Q3 as all ontologies have their metadata described in an external file (AgroPortal's metadata record) and because AgroPortal offers a bidirectional link between an ontology and its metadata.[43] An implementation of O'FAIRe outside of AgroPortal would have required a metadata property to identify and locate explicitly the metadata.

When assessing F4, we observed:

- 51 ontologies (34%) are registered at least in two libraries, mostly the OBO Foundry, Agrisemantics Map of Standards and FAIRSharing; 46 ontologies (30%) are registered in only one library (Q1).
- Only 2 ontologies (1%) are declared available in at least five open ontology repositories; 24 ontologies (16%) are registered at least in four repositories; 11 (7%) in at least in three; 36 (24%) in at least two; and finally, 76 (51%) are registered in only one repository (Q2).

Fig. 7. Average normalized *FAIRSubPrincipleScore* for F.

We observe an average normalized score of 90 for A, detailed for each sub-principle in Fig. 8. When assessing A1, A1.1 and A1.2 we observed:

- 70 ontologies (46%) have at least two resolvable identifiers (A1.Q1). A1.Q2 is directly addressed by the repository which provides an external metadata URI which resolves to the metadata record.
- 58 ontologies (38%) support at least content negotiation with 6 formats; and 86 (57%) support at least 5 formats (A1.Q3). These numbers include 4 formats systematically supported by AgroPortal for metadata.
- 142 ontologies (95%) declare being accessible through a SPARQL endpoint (A1.Q4). Indeed, this field is either provided by the ontology developer or provided by AgroPortal's SPARQL endpoint.
- 100% of ontologies are accessible using an open, free, and universal communication protocol: HTTP. This criterion is independent of the ontologies and depends on AgroPortal (Q1, Q2, and Q3 in A1.1).
- Similarly, 100% of ontologies are accessible via authentication and authorization as this feature is supported by AgroPortal (A1.2).

When assessing A2, we observed all ontologies are totally compliant with the A2 principle because AgroPortal supports versioning (Q1, Q2, and Q3). Plus, 144 ontologies (96%) clearly inform about their status.

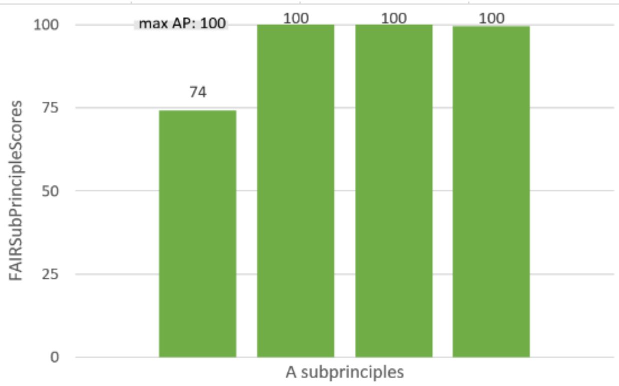


Fig. 8. Average normalized *FAIRSubPrincipleScore* for A.

We observe an average normalized score of 42 for I, detailed for each sub-principle in Fig. 9. When assessing I1, we observed:

- Only 4 ontologies (2%) have not explicitly declared their format (Q1 and Q2).
- 126 ontologies (85%) inform about their syntax (Q3). 130 (90%) describe also the formality level (Q4). These two relatively high scores are due to the fact that the corresponding properties are closely curated by AgroPoral.
- In contrast, only 2 ontologies (1%) inform about their availability in other formats (Q5).

As depicted in Fig. 9, I2 and I3 are hard to meet as the scores are impacted by the questions not yet implemented in O'FAIRe (i.e., I2.Q3, I2.Q5, I3.Q1, I3.Q2). When assessing I2, we observed:

- 149 ontologies (100%) import other FAIR vocabularies (I2.Q1), reuse terms for other vocabularies (I2.Q2), and use valid URIs to encode some metadata values (I3.Q3).
- 38 ontologies (25%) are aligned to other vocabularies (I2.Q4).
- 120 ontologies (80%) provide information about the relation to or influence of other vocabularies (I2.Q6).

In I3, we focused on the assessment of the included references to other (meta)data-based without dealing with the qualification aspect; in our case, a "qualified reference" meant an additional link that would help in understanding the content or the provenance of any ontologies; here we consider projects and endorsing organizations.

Finally, we observe an average normalized score of 38 for R, detailed for each sub-principle in Fig. 10. When assessing R1, we observe:

- 114 ontologies (76%) provide at least 4 information about how classes or concepts are defined (Q1) and 133 ontologies (89%) provide at least one information about their hierarchies (Q2). These relatively high scores are explained by the fact that 8 properties are enforced

[43] In AgroPortal, an "ontology" is linked to several "submissions". Most of the metadata is attached to the submission object.

automatically by AgroPortal which uses them in its internal model. If the information is not entered by the ontology developer, one is assigned by default; however, we have observed a better curation of these properties is required as default properties are not necessarily valid.

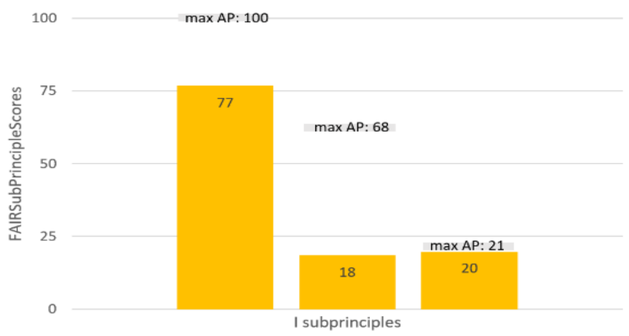


Fig. 9. Average normalized *FAIRSubPrincipleScore* for I.

When assessing R1.1, we observe:

- 91 ontologies (61%) provide clear licensing information with a resolvable license URI with content negotiation (Q1). Although this field is frequently included in ontologies or in their documentations, the AgroPortal team works on harmonizing reference to licenses usually using URIs from the CC vocabulary.
- AgroPortal systematically distinguishes between two cases: public or private ontologies. Therefore, in our prototype, we consider 100% ontologies document their access rights thanks to AgroPortal (Q2).
- 140 ontologies (93%) documents at least one information about usage guidelines and copyrights holder (Q3).

When assessing R1.2, we observe:

- 115 ontologies (77%) provide at least one information about the actors (creator, contributor, curator, translator) involved in the development (Q1). And 137 (87%) provide at least one information about general provenance (Q2).
- In contrast, only 6 ontologies (4%) provide information about accrual methods and policies (Q3).
- A good versioning for 137 ontologies (91%) with the description of their current version and links between their previous versions (Q4). Those aspects are also enforced by the repository.
- 76 ontologies (51%) inform about the tools and methods used in the development (Q6); 66 ontologies (44%) inform about their funding organization (Q8). But, only 10 ontologies (6%) describe their rationale e.g., competency questions and related tasks (Q7).

When assessing R1.3, we observe:

- A good description of endorsing organizations or related projects in 123 ontologies (82%) information are available (Q1).
- 136 ontologies (92%) are open (Q3) among them 92 ontologies (61%) are included in a specific community set or group (Q2).

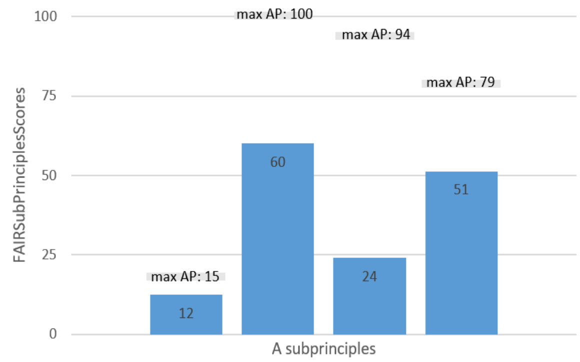


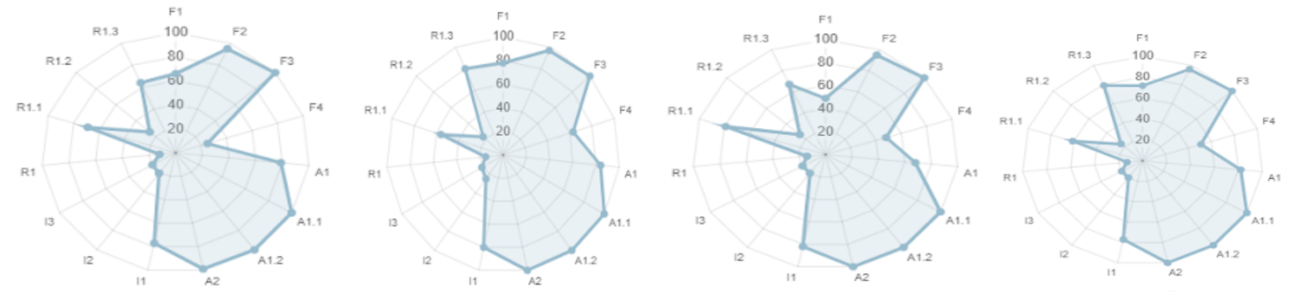
Fig. 10. Average normalized *FAIRSubPrincipleScore* for R.

Fig. 11 describes the average normalized FAIR scores for each studied group of ontologies in AgroPortal namely, INRAE, OBO-F, CROP, and WHEAT and Table 2 shows statistics. Hereafter, we highlight some findings:

- Scores (Fig. 11) are similar for all sub-principles except in F1, F2, F4, A1, R1.1 and R1.3.
- Average FAIR scores are relatively the same (Table 2). It is clear that fulfilling the FAIR principles is challenging for all communities, even the OBO-F that has been doing a considerable effort in ontology and metadata maintenance [53].
- Ontologies in the OBO-F group have the higher F1 scores (identifiers) thanks to a rigorous use of PURLs; However, the CROP group has a low score because 99% of ontologies are without URIs.
- Ontologies in OBO-F group adopt a relatively "good" usage of version URIs compared to other groups; 65% ontologies have a resolvable version URIs but only 40% in WHEAT, 1% in INRAE, and 0% in CROP.
- Ontologies in the WHEAT group follow "good" practices in reusability in particular, aspects related to license, access rights, and funding organization.
- All ontologies in the CROP group have resolvable licenses. We notice also that INRAE scores are also good because 27 ontologies (84%) are fully compliant with R1.1.Q1 (license). This reflects efforts from these communities to clarify licensing.
- 72% of CROP ontologies describe their funding organization (R1.2 Q8), but only 58% in INRAE, 40% in WHEAT, and 26% in OBO-F.
- The majority of OBO-F ontologies (99%) do not document well the permissions, usage guidelines and copyrights holder (R1.1.Q3).
- Best ontology in INRAE is Agriculture Experiments Ontology (AEO), in CROP is Banana Ontology (CO_325), and in WHEAT is Environment Ontology.

Table 2. FAIR scores statistics for the 4 studied groups.

	INRAE n=31	OBO-F n=24	CROP n=37	WHEAT n=20
Min	53	51	25	44
Max	67	71	67	70
Average	61	65	59	63
Median	61	67	60	63



(a) INRAE FAIRness (b) OBO-F FAIRness (c) CROP FAIRness (d) WHEAT FAIRness

Fig. 11. FAIRness assessment for each studied group of semantic resources in AgroPortal.

5 Discussion and perspectives

5.1 About FAIRness assessment and O'FAIRe

Our projection of the 15 FAIR principles for ontology and ontology metadata shows they are easy to interpret at first sight because they are described with short sentences and a clear terminology. However, digging in more details, we can say only 8 sub-principles are simple to put into practice and evaluate; namely, F1, F3, F4, A1, A1.1, A1.2, I1, R1.1; the rest are not obvious and are subjective to some interpretation:

- F2 deserves a specific treatment as it is about metadata –in general– but many other sub-principles are also about a specific metadata aspect. Possibly, anything describing an ontology that does not explicitly fit in another principle can be used, somehow, to assess F2. In our methodology, the upper bound is set by properties included in MOD1.4.
- F2 and R1 are somehow overlapping, with "rich metadata" in F2 and "richly described with a plurality of (...) attributes" in R1. The difference between them could be clearly specified. We have decided to focus R1 on metadata related to explicitly describing the content of ontologies (labels, definitions, hierarchy, etc.).
- A2 requires clear ontology and metadata preservation strategies; these strategies are not trivial to verify/assess often by lack of historical or versioning information.
- I2 focuses on the use of FAIR vocabularies, which inherently creates a recursive situation when assessing the level of FAIRness of the reused vocabularies; this is not really addressed in our methodology, but a perspective would be to weigh the score of I2 questions with a FAIR Score obtained by our own methodology when available.
- I3 is a little vague or too wide as it relates to all the possible "relations" between ontologies and data; our choice to evaluate it with cross-references is certainly limited.
- R1.2 is very hard to evaluate as provenance can be understood with a very large spectrum. An approach like ours based on metadata descriptions is limited by the availability of metadata properties to describe metadata provenance. MOD1.4 includes 6 explicit provenance properties but others could be considered and included in the methodology.
- R1.3 is community-specific and thus deserves a community-specific-subjective interpretation and evaluation; what is a standard in the biomedical domain might not be relevant in agronomy or industry, etc. Some communities have established their guidelines, others not yet.

5.2 About O'FAIRe questions

A key feature of our approach is that the final credits are are independent of the number of questions. FAIRness assessment questions are the backbone of O'FAIRe methodology; they can be adapted or completed to refine the evaluation of one specific aspect without changing the overall methodology. For example, if a fourth question related to a sub-principle is added, then the total number of credits assigned for this sub-principle will be re-dispatched on the four questions rather than on three, then refining the assessment process without changing the overall balance of principles relatively to each other. The current list of questions are derived both from our experience working on ontology repository and services for years [39] but also our participation –and sometime leading roles– in interest groups such as EUDAT Semantics, RDA VSSIG, GO-FAIR INs, FAIRsFAIR workshops [5]. Many of O'FAIRe questions are generic –i.e., not specific to ontologies– and could be reused by other communities for implementing FAIRness assessment methods for other kinds of digital objects. Of course, there are a few limits in our methodology:

- Due to a few missing metadata properties in MOD and some limits in AgroPortal the current O'FAIRe prototype does not implement 11 questions. MOD is currently being revised and a v2 –based on the DCAT W3C Recommendation– shall address our needs with respect to missing properties. AgroPortal technical limitations will be addressed along the way to improve the system.[44]

[44] For example, we do not yet have a mechanism to evaluate the quantities (i.e., the ratio over the size of the ontology) and qualities of mappings (I2) or qualified references (I3) from an ontology. As another example, in some cases, O'FAIRe just checks the existence of a value for a metadata property but not the exactitude or relevance of that value.

- O'FAIRe approach to rely on metadata properties to evaluate a question means that it is bounded to a metadata model (such as MOD1.4) that must offer the required properties. It is dependent on either: (i) the population of the metadata by the developer or curator; (ii) other components to compute automatically the value of a metadata property such as, the component computing metrics in AgroPortal.
- The integrated FAIRness assessment grid which fixes the importance and balance of each FAIR principles and sub-principles shall be updated to include recent or future relevant works e.g., as proposed by Cox et al. [10], or the final iteration of the FAIRsFAIR recommendations (to come in 2022). Note, eventually the guidelines and recommendations shall certainly converge and new entries will mathematically less influence the grid. Therefore, we do not expect the grid to change much after a certain time.
- O'FAIRe v2's 61 questions would need a community evaluation now that a prototype is available and demonstrable. Our previous tentative discussions of these questions before their implementation turned out unproductive because they were too detached from reality. In the future, it would be useful to collect community feedback to consolidate and validate our proposal and avoid biases that certainly come from our interpretation of how to make semantic resources FAIR. For instance, Annex 4 and Annex 5 contains subjective coarse-grained appreciations that in the future we would called for vote in the context of RDA VSSIG.

5.3 About O'FAIRe implementation and results

With the AgroPortal repository, our results show interoperability and reusability aspects are somehow less addressed. Comparatively, accessibility is quite good, intrinsically because semantic resources are digital objects, mostly relying on semantic Web technologies which address (with HTTP and URIs) some of the FAIR accessibility requirements; in our implementation of O'FAIRe in AgroPortal's accessibility is also taken care by the repository.

O'FAIRe implementation presented Section 4.1.1 differs from existing initiatives, as it proposes an automatic FAIRness assessment and offers suggestions for improving the obtained scores. It is compatible with any ontology repository based on the OntoPortal technology (originally developed for the NCBO BioPortal) and implementing the MOD1.4 ontology metadata model (or an equivalent one offering the same coverage with metadata properties). Currently, it can be used on AgroPortal, previously presented, and the SIFR BioPortal, a French biomedical terminologies repository (http://bioportal.lirmm.fr) [54] built to implement the SIFR Annotator ontology-based text annotation tool [55]. Collaborations within the OntoPortal Alliance will enable us to extend and maybe customize O'FAIRe for other repositories such as the NCBO BioPortal, LifeWatch's EcoPortal or Fraunhofer's MatPortal. Eventually, we can also refactor our tool to evaluate the level of FAIRness of an ontology simply based on the ontology file –and its included or referenced metadata. But we believe that today, such an approach would be unfruitful due to the lack of metadata descriptions present in ontologies taken into the wild [8].

One limitation of our analysis is that the current FAIRness assessment scores are calculated on the basis of the current state of AgroPortal's ontologies descriptions. Indeed, the scores will necessarily change in the near future with the release of O'FAIRe in AgroPortal (Feb. 2022) and its usability for ontology developers to improve the FAIR score of their resources. AgroPortal team will also necessarily use this information in the future metadata curation process. This situation is in fact, desirable as the overall objective of O'FAIRe is indeed to augment the adoption of the FAIR principles for ontologies and semantic resources. In other words, it means that the impact of O'FAIRe will be measurable in a few months after it has been used to increase the level of FAIRness.

6 Conclusion

FAIR is a journey, not necessarily a destination; a key element to become FAIR is to use FAIR vocabularies and ontologies. But, how do we know if an ontology is FAIR? The need to construct a common framework to define ontology FAIRness and harmonize multiple existing assessment guidelines and methods (generic and specific) was clearly identified by the RDA Vocabulary and Semantic Services Interest Group and the H2020 FAIRsFAIR actions. In this paper, we guide the semantic community to put the FAIR principles into practice and enable them to qualify the *degree of FAIRness* of their semantic resource. Our work addresses several scientific and technical challenges regarding the implementation of the 15 FAIR principles for ontologies and semantic resources. It provides solutions that would facilitate the adoption of FAIR. Based on preliminary results, we analyzed how easy or hard it is for semantic resources to adhere to the principles; analysis led us to many conclusions and perspectives for FAIR ontologies.

Author contributions

EA compiled related work. EA and CJ conceived the methodology and jointly wrote the article. EA and SB implemented O'FAIRe prototype and integration in AgroPortal. CJ envisioned and led the research action.

Acknowledgments

This work has been supported by the *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). We also thank the VSSIG (Vocabulary and Semantic Services Interest Group) of the Research Data Alliance and the H2020 FAIRsFAIR project T2.2 on "FAIR Semantics" for fruitful discussions and exchanges.

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] B. Mons, *Data Stewardship for Open Science*. Chapman and Hall/CRC, 2018. doi: 10.1201/9781315380711.

[3] M. D. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. Bonino da Silva Santos, and M. Dumontier, "A design framework and exemplar metrics for FAIRness," *Scientific Data*, vol. 5, p. 180118, Jun. 2018, doi: 10.1038/sdata.2018.118.

[4] M. D. Wilkinson *et al.*, "Evaluating FAIR maturity through a scalable, automated, community-governed framework," *Scientific Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1038/S41597-019-0184-5.

[5] 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.

[6] A. Devaraju *et al.*, "From conceptualization to implementation: Fair assessment of research data objects," *Data Science Journal*, vol. 20, no. 1, pp. 1–14, Feb. 2021, doi: 10.5334/DSJ-2021-004.

[7] C. Jonquet, "FAIR data requires FAIR ontologies, how do we do?," *F1000Res*, vol. 7, Mar. 2018, doi: 10.7490/F1000RESEARCH.1115343.1.

[8] 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.

[9] C. Jonquet *et al.*, "AgroPortal: A vocabulary and ontology repository for agronomy," *Computers and Electronics in Agriculture*, vol. 144, pp. 126–143, 2018, doi: 10.1016/j.compag.2017.10.012.

[10] 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.

[11] 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*, G. Cota, M. Daquino, and G. L. Pozzato, Eds. IOS Press, 2020. doi: 10.3233/SSW200034.

[12] 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, doi: 10.1371/journal.pcbi.1004743.

[13] E. Amdouni and C. Jonquet, "FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies," Nov. 2021. doi: 10.13140/RG.2.2.23265.84321.

[14] Y. Ding and D. Fensel, "Ontology Library Systems: The key to successful Ontology Re-use," in *1st Semantic Web Working Symposium, SWWS'01*, 2001, pp. 93–112.

[15] M. D'Aquin and N. F. Noy, "Where to Publish and Find Ontologies? A Survey of Ontology Libraries.," *Web semantics*, vol. 11, no. August, pp. 96–111, Mar. 2012, doi: 10.1016/j.websem.2011.08.005.

[16] A. Toulet, B. Dutta, and C. Jonquet, "Assessing the Practice of Ontology Metadata: A Survey Result," Montpellier, France, 2018. doi: 10.5281/zenodo.3484530.

[17] 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.

[18] B. Mons, C. Neylon, J. Velterop, and M. Dumontier, "Cloudy, increasingly FAIR ; revisiting the FAIR Data guiding principles for the European Open Science Cloud," vol. 37, pp. 49–56, 2017, doi: 10.3233/ISU-170824.

[19] Directorate-General for Research and Innovation, "Turning FAIR into reality -," 2018. doi: 10.2777/54599.

[20] 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.

[21] U. Schwardmann, "Digital objects – FAIR digital objects: Which services are required?," *Data Science Journal*, vol. 19, no. 1, pp. 1–6, Apr. 2020, doi: 10.5334/DSJ-2020-015.

[22] S. Cox and J. Yu, "OzNome 5-star Tool: A Rating System for making data FAIR and Trustable – 2022 eResearch Australasia Conference,"

[23] 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, doi: 10.1016/J.CELS.2019.09.011.

[24] 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.

[25] 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. doi: 10.5281/ZENODO.5914307.

[26] C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data - The Story So Far," *Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009, doi: 10.4018/jswis.2009081901.

[27] K. Janowicz, P. Hitzler, B. Adams, D. Kolas, and C. Vardeman, "Five stars of Linked Data vocabulary use," *Semantic Web*, vol. 5, no. 3, pp. 173–176, Jan. 2014, doi: 10.3233/SW-140135.

[28] 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.

[29] 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.

[30] 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.

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

[32] M. L. Zeng and J. Clunis, "FAIR + FIT: Guiding Principles and Functional Metrics for Linked Open Data (LOD) KOS Products," *Journal of Data and Information Science*, vol. 5, no. 1, pp. 93–118, Feb. 2020, doi: 10.2478/JDIS-2020-0008.

[33] S. J. D. Coxid, A. N. Gonzalez-Beltranid, B. Magagna, and M.-C. Marinescu, "Ten simple rules for making a vocabulary FAIR," 2021, doi: 10.1371/journal.pcbi.1009041.

[34] J. Frey, D. Streitmatter, F. Götz, S. Hellmann, and N. Arndt, "DBpedia Archivo: A Web-Scale Interface for Ontology Archiving Under Consumer-Oriented Aspects," in *International Conference on Semantic Systems, SEMANTICS'20*, Sep. 2020, vol. 12378 LNCS, pp. 19–35. doi: 10.1007/978-3-030-59833-4_2.

[35] C. Frasson and G. Kostopoulos, Eds., *Brain Function Assessment in Learning*, vol. 10512. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-67615-9.

[36] Suarez-Figueroa, J. Hartmann, Y. Sure, P. Haase, and M. Suarez-Figueroa, “OMV–ontology metadata vocabulary,” in *Workshop on Ontology Patterns for the Semantic Web, WOP’05*, 2005, p. 9.
[37] P.-Y. Vandenbussche, G. A. Atemezing, M. Poveda-Villalon, and B. Vatant, “Linked Open Vocabularies (LOV): a gateway to reusable semantic vocabularies on the Web,” *Semantic Web*, vol. 1, pp. 1–5, 2014, doi: 10.3233/SW-160213.
[38] N. F. Noy *et al.*, “BioPortal: ontologies and integrated data resources at the click of a mouse,” *Nucleic Acids Research*, vol. 37, no. (web server), pp. 170–173, 2009, doi: 10.1093/nar/gkp440.
[39] R. G. Côté, P. Jones, R. Apweiler, and H. Hermjakob, “The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries.,” *BMC Bioinformatics*, vol. 7, p. 97, Jan. 2006, doi: 10.1186/1471-2105-7-97.
[40] N. Fiore, B. Magagna, and D. Goldfarb, “EcoPortal: A Proposition for a Semantic Repository Dedicated to Ecology and Biodiversity,” in *2nd International Workshop on Semantics for Biodiversity, S4BioDiv’17*, 2017.
[41] J. Hartmann, R. Palma, and A. Gómez-Pérez, “Ontology Repositories,” *Handbook on Ontologies*, pp. 551–571, 2009, doi: 10.1007/978-3-540-92673-3.
[42] C. Jonquet, “Ontology Repository and Ontology-Based Services – Challenges, contributions and applications to biomedicine & agronomy,” 2019. HDR, University of Montpellier.
[43] J. Graybeal, C. Jonquet, N. Fiore, and M. A. Musen, “Adoption of BioPortal’s Ontology Registry Software: The Emerging OntoPortal Community,” 2019.
[44] A. Laadhar, E. Abrahão, and C. Jonquet, “Analysis of Term Reuse, Term Overlap and Extracted Mappings Across AgroPortal Semantic Resources,” in *22nd International Conference on Knowledge Engineering and Knowledge Management, EKAW’20*, Sep. 2020, vol. 12387 LNAI, pp. 71–87. doi: 10.1007/978-3-030-61244-3_5.
[45] M. Courtot *et al.*, “MIREOT: The minimum information to reference an external ontology term,” *Applied Ontology*, vol. 6, no. 1, pp. 23–33, 2011, doi: 10.3233/AO-2011-0087.
[46] A. Laadhar, É. Abrahão, and C. Jonquet, “Investigating One Million XRefs in Thirthy Ontologies from the OBO World,” in *11th International Conference on Biomedical Ontologies, ICBO’20*, Sep. 2020
[47] R. C. Jackson, J. P. Balhoff, E. Douglass, N. L. Harris, C. J. Mungall, and J. A. Overton, “ROBOT: A Tool for Automating Ontology Workflows,” *BMC Bioinformatics*, vol. 20, 1, pp. 1–10, Jul. 2019, doi: 10.1186/S12859-019-3002-3.
[48] A. Gangemi, “Ontology Design Patterns for Semantic Web Content,” in *4th International Semantic Web Conference, ISWC 2005*, 2005, vol. 3729, pp. 262–276. doi: 10.1007/11574620_21.
[49] B. Smith *et al.*, “The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration,” *Nature Biotechnology*, vol. 25, no. 11, pp. 1251–1255, 2007, doi: 10.1038/nbt1346.
[50] L. Cooper *et al.*, “The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics.,” *Nucleic Acids Res*, vol. 46, no. D1, pp. D1168–D1180, Jan. 2018, doi: 10.1093/nar/gkx1152.
[51] R. Shrestha *et al.*, “Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice,” *Frontiers in Physiology*, vol. 3 AUG, no. August, pp. 1–10, 2012, doi: 10.3389/fphys.2012.00326.
[52] E. D. Yeumo *et al.*, “Developing data interoperability through standards: a wheat community use case,” *F1000 Research*, vol. 6, no. 1843, 2017, doi: 10.12688/f1000research.12234.1.
[53] R. Jackson *et al.*, “OBO Foundry in 2021: operationalizing open data principles to evaluate ontologies,” *Database*, vol. 2021, no. 0, pp. 1–9, Sep. 2021, doi: 10.1093/DATABASE/BAAB069.
[54] C. Jonquet, A. Annane, K. Bouarech, V. Emonet, and S. Melzi, “SIFR BioPortal : Un portail ouvert et générique d’ontologies et de terminologies biomédicales françaises au service de l’annotation sémantique,” 2016.
[55] A. Tchechmedjiev, A. Abdaoui, V. Emonet, S. Zevio, and C. Jonquet, “SIFR Annotator: Ontology-Based Semantic Annotation of French Biomedical Text and Clinical Notes,” *BMC Bioinformatics*, vol. 19, pp. 405–431, 2018, doi: 10.1186/s12859-018-2429-2.
[56] M. L. Zeng, “Knowledge Organization Systems (KOS),” *Knowledge organization*, vol. 35, no. 2–3, pp. 160–182, 2008.
[57] C. Caracciolo *et al.*, “39 hints to facilitate the use of semantics for data on agriculture and nutrition,” *Data Science Journal*. 19(1), p.47. doi: 10.5334/dsj-2020-047.
[58] E. Schultes, B. Magagna, K. M. Hettne, R. Pergl, M. Suchánek, and T. Kuhn, “Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence,” International Conference on Conceptual Modeling, Workshops vol. 12584 LNCS, pp. 138–147, 2020, doi: 10.1007/978-3-030-65847-2_13/FIGURES/2.

Supplementary material

Annex 1. Integrated FAIRness assessment grid for semantic resources and ontologies.

Principle		Baseline	SHARC	FDMM	5-stars V	MIRO	FAIRsFAIR	Poveda et al.	Credits	
F	**F1**	10	8	12	0	6	2	3	41	**113**
	F2	10	8	2	0	5	1	1	27	
	F3	10	8	2	0	0	1	0	21	
	F4	10	8	3	0	0	2	1	24	
A	**A1**	10	6	18	3	3	1	2	43	**113**
	A1.1	10	6	11	0	0	1	0	28	
	A1.2	10	6	5	0	0	1	0	22	
	A2	10	6	3	0	0	1	0	20	
I	**I1**	10	4	12	1	12	3	2	44	**109**
	I2	10	4	7	0	9	2	0	32	
	I3	10	4	12	1	3	2	1	33	
R	**R1**	10	9	3	1	6	0	3	32	**143**
	R1.1	10	9	12	0	3	2	1	37	
	R1.2	10	9	3	0	12	3	1	38	
	R1.3	10	9	12	0	0	3	2	36	
Total credits										**478**

Annex 2. List of 71 "extra" metadata properties from MOD 1.4 and their alignment with MIRO guidelines.

Must	omv:acronym, dct:title, dct:alternative, skos:hiddenLabel, dct:description, foaf:page, omv:resourceLocator, omv:keywords, dct:coverage, foaf:homepage, vann:example, vann:preferredNamespaceUri, void:uriRegexPattern, idot:exampleIdentifier, dct:publisher, dct:subject, owl:backwardCompatibleWith, door:comesFromTheSameDomain, omv:knownUsage, dct:audience, doap:repository, doap:bugDatabase, doap:mailing-list, mod:hasEvaluation
Should	mod:metrics, omv:numberOfClasses, omv:numberOfIndividuals, omv:numberOfProperties, mod:numberOfDataProperties, mod:numberOfObjectProperties, omv:numberOfAxioms, mod:numberOfLabels, mod:byteSize
Optional	vann:preferredNamespacePrefix
No mapping	dct:language, dct:abstract, mod:analytics, dct:bibliographicCitation, rdfs:comment, foaf:depiction, foaf:logo, voaf:toDoList, schema:award, schema:associatedMedia, owl:incompatibleWith, dct:hasPart, schema:workTranslation, door:hasDisparateModelling, voaf:usedBy, voaf:hasDisjunctionsWith, omv:keyClasses, void:rootResource, mod:browsingUI, mod:sampleQueries, void:propertyPartition, void:classPartition, void:dataDump, void:openSearchDescription, void:uriLookupEndpoint, schema:comment, dct:created, dct:modified, dct:valid, dct:dateSubmitted, pav:curatedOn, omv:isOfType, mod:classesWithMoreThan25Children, mod:classesWithOneChild, mod:classesWithNoDefinition, mod:maxChildCount

Annex 3. List of selected libraries and repositories taken into consideration for the evaluation of the F4 questions. The star symbol (*) distinguishes the ones supporting authentication mechanisms. The list is open for modification/suggestion on GitHub: https://github.com/agroportal/fairness/blob/master/src/main/resources/config/common/catalogs.config.json

	Name	**URL**
Library	OBO Foundry	http://www.obofoundry.org/
	CROP Ontology Curation Tool	http://www.cropontology.org/
	Agrisemantics	http://vest.agrisemantics.org/
	FAIRSharing	http://fairsharing.org/
	ONKI Ontology Library Serv.	http://onki.fi/
	MMI ORR	http://mmisw.org/
	ROMULUS	http://thezfiles.co.za/ROMULUS/
	DAML Ontology Library	http://www.daml.org/
	Colore	http://stl.mie.utoronto.ca/colore/
	BARTOC	http://bartoc.org/
	TaxoBank	http://www.taxobank.org/
	TaxoBank	http://www.taxobank.org/
	LusTRE	http://linkeddata.ge.imati.cnr.it/
	LOV4IoT	http://lov4iot.appspot.com/
	VOCAB OEG	http://vocab.linkeddata.es/
	LiveSchema	http://liveschema.eu/
	Protege Ontology library	http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library
Repository	BioPortal *	http://bioportal.bioontology.org/
	AgroPortal *	http://agroportal.lirmm.fr/
	SIFR BioPortal *	http://bioportal.lirmm.fr/
	MedPortal *	http://medportal.bmicc.cn/
	LifeWatch EcoPortal *	http://ecoportal.lifewatchitaly.eu/
	MatPortal *	http://matportal.org/
	IndustryPortal	http://industryportal.enit.fr/
	EBI Ontology Lookup Serv.	https://www.ebi.ac.uk/ols/
	Ontobee	http://www.ontobee.org/
	LOV	http://lov.linkeddata.es/dataset/lov
	AberOwl	http://aber-owl.net/
	OntoHub	http://ontohub.org/
	CISMeF HeTOP*	http://www.hetop.eu/hetop/
	FINTO	https://finto.fi/en/
	ANDC RVA	https://vocabs.ardc.edu.au/
	NERC NVS	https://vocab.nerc.ac.uk/
	GFBIO IS	http://terminologies.gfbio.org/
	LOTERRE	https://www.loterre.fr
	Planteome	http://www.planteome.org/
	CALIPER	http://datalab.review.fao.org/datalab/caliper/web/

Annex 4. O'FAIRe proposed classification of knowledge representation languages with evaluation (1 to 5 stars) of each characteristic required by sub-principle I1. Note: our interpretation of each characteristic is: Formal: How easy is it for machines to understand the semantics of the knowledge encoded in a resource? Accessible: How easy is it for machines to access and process the content of the resource? Shared: Is the representation language used widely adopted and shared? Is it supported by a well-known community? Broadly applicable: How much the language can be used for a wide range of ontologies and metadata representation independently of the task or the domain?

	Formal	**Accessible**	**Shared**	**Broadly applicable**	**Total points**
OWL	*****	*****	***	*****	18
OBO	****	****	*	**	11

RDFS	**	*****	**	***	12
SKOS	***	*****	***	****	15
CSV	*	**	****	**	9
XML	*	***	****	**	10
PDF	-	-	****	*	5
TXT	-	*	*****	*	7

Annex 5. O'FAIRe selected metadata vocabularies and their proposed evaluation with a coarse grain "4-level based FAIRness coefficient" manually assigned (from 1 to 4 stars); recognized by I2Q7. Available also on GitHub: https://github.com/agroportal/fairness/blob/master/src/main/resources/config/common/metadata.voc.config.json

Prefix	Name	URL	Coef
rdfs	RDF Schema	https://www.w3.org/TR/rdf-schema/	***
owl	OWL 2 Web Ontology Language	https://www.w3.org/TR/owl2-quick-reference/	***
skos	Simple Knowledge Organization System	https://www.w3.org/TR/skos-reference/	***
dc	Dublin Core (DCMI Metadata Terms)	http://purl.org/dc/elements/1.1/	***
dct terms	Dublin Core (DCMI Metadata Terms)	http://purl.org/dc/terms/	***
omv	Ontology Metadata Vocabulary	http://omv2.sourceforge.net/	**
mod	Metadata for Ontology Description and Publication	http://www.isibang.ac.in/ns/mod.html	**
door	Descriptive Ontology of Ontology Relations	http://oro.open.ac.uk/24326/1/keod09.pdf	*
voaf	Vocabulary of a Friend	http://purl.org/vocommons/voaf#	**
void	Vocabulary of Interlinked Datasets	https://www.w3.org/TR/void/	**
idot	identifiers.org Terms	http://identifiers.org/idot	*
vann	Vocabulary for annotating vocabulary descriptions	http://purl.org/vocab/vann/	**
dcat	Data Catalog Vocabulary	https://www.w3.org/TR/vocab-dcat/	***
adms	Asset Description Metadata schema	https://www.w3.org/TR/vocab-adms/	**
schema	schema.org	http://schema.org/Dataset	***
foaf	Friend of a Friend Vocabulary	http://xmlns.com/foaf/spec/	***
doap	Description of a Project	https://github.com/edumbill/doap/wiki	**
cc	Creative Commons Rights Expression Language	http://creativecommons.org/ns	***
prov	Provenance Ontology	https://www.w3.org/TR/prov-o/	***
oboInowl	OboInOwl Mappings	http://www.geneontology.org/formats/oboInOwl#	**
sd	SPARQL 1.1 Service Description	https://www.w3.org/TR/sparql11-service-description/	*
nkos	Networked Knowledge Organization Systems Dublin Core Application Profile (NKOS AP)	http://nkos.slis.kent.edu/nkos-ap.html	*
cito	The citation Typing Ontology	http://purl.org/spar/cito/	*
wdrs	Protocol for Web Description Resources (POWDER): POWDER-S Vocabulary (WDRS)	https://www.w3.org/2007/05/powder-s	**
bibo	The Bibliographic Ontology	http://lov.okfn.org/dataset/lov/vocabs/bibo	**

disco	DDI-RDF Discovery Vocabulary	http://rdf-vocabulary.ddialliance.org/discovery.html	**
oa	Open Annotation Data Model	https://www.w3.org/ns/oa	***
iao	Information Artifact Ontology	http://purl.bioontology.org/ontology/IAO	**
sio	Semanticscience Integrated Ontology (SIO)	http://semanticscience.org/resource/	**
void-ext	Extensions to the Vocabulary of Interlinked Datasets (VoID)	http://ldf.fi/schema/void-ext/	*
ro	Relations Ontology		*
swoogle	Swoogle Ontology	http://daml.umbc.edu/ontologies/webofbelief/1.4/swoogle.owl#	**
edam	EDAM	http://purl.obolibrary.org/obo/edam#	**