



**HAL**  
open science

# Une méthodologie et un outil d'évaluation du niveau de "FAIRness" pour les ressources sémantiques : le cas d'AgroPortal

Emna Amdouni, Clement Jonquet

## ► To cite this version:

Emna Amdouni, Clement Jonquet. Une méthodologie et un outil d'évaluation du niveau de "FAIRness" pour les ressources sémantiques : le cas d'AgroPortal. IC 2021 - 32es Journées francophones d'Ingénierie des Connaissances, Jun 2021, Bordeaux, France. pp.11-19. lirmm-03232615

**HAL Id: lirmm-03232615**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-03232615>**

Submitted on 21 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Une méthodologie et un outil d'évaluation du niveau de "FAIRness" pour les ressources sémantiques : le cas d'AgroPortal

Emna Amdouni<sup>1</sup> et Clément Jonquet<sup>1,2</sup>

<sup>1</sup> LIRMM, Univ. de Montpellier, CNRS, France

<sup>2</sup> MISTEA, Univ. de Montpellier, INRAE, Institut Agro, France

[emna.amdouni@lirmm.fr](mailto:emna.amdouni@lirmm.fr) et [jonquet@lirmm.fr](mailto:jonquet@lirmm.fr)

## Résumé

*Les principes "FAIR" définissent un ensemble de caractéristiques que les données et leurs métadonnées devraient présenter pour être Faciles à trouver, Accessibles, Interopérables et Réutilisables. Également, suivant le principe I2, une ontologie, et plus généralement une ressource sémantique, devrait aussi être "FAIR". Des critères spécifiques aux ontologies commencent à apparaître, mais il n'existe toujours pas de mécanismes pour évaluer le degré de mise en œuvre de ces principes. Nous proposons une approche d'évaluation automatique du niveau de "FAIRness" d'une ontologie qui se base sur une représentation riche et structurée des métadonnées. Nous nous reposons sur le modèle de métadonnées MOD et avons développé un prototype pour AgroPortal, un portail de ressources sémantiques et d'ontologies en agronomie.*

## Mots-clés

*Science ouverte, principes FAIR, FAIRness, ontologies, ressources sémantiques, portail d'ontologies, AgroPortal*

## Abstract

*The "FAIR" principles define a set of characteristics data and their metadata should respect in order to be Findable, Accessible, Interoperable, and Reusable. Per principle I2, an ontology, and more generally a semantic resource, should also be FAIR. Ontology-specific criteria begin to emerge, but there is still no mechanism to assess the degree of implementation of these principles. We propose an automatic evaluation method of the level of "FAIRness" of an ontology which is based on a rich and structured representation of the metadata. We rely on the MOD metadata model and have developed a prototype for AgroPortal, a repository of semantic resources and ontologies in agronomy.*

## Keywords

*Open science, FAIR principles, FAIRness, ontologies, semantic resources, ontology repository, AgroPortal*

## 1 Introduction

En 2014, un groupe de chercheurs et d'éditeurs (appelé FORCE 11) a établi des principes de création et de partage des données scientifiques afin de favoriser leur exploitation et leur réutilisation d'une manière automatique entre autres via le Web. Ces principes sont introduits sous le nom de FAIR (acronyme de Findable, Accessible, Interoperable et Reusable). Chacun

des quatre principes FAIR décrit un ensemble de caractéristiques génériques que les données (et leurs métadonnées) devraient avoir mais ils ne précisent pas, ni ne préconisent, les mécanismes de mise en œuvre [1], [2].

Ci-après, nous présentons brièvement la signification de chaque aspect F-A-I-R : Premièrement, nous parlons de données *faciles à trouver* lorsqu'elles sont suffisamment décrites avec des métadonnées et hébergées ou indexées dans une librairie ou un portail accessible ouvertement. La mise en œuvre de ce principe implique que les données, métadonnées et autres ressources doivent avoir un identifiant unique et persistant qui les rend trouvables et référençables par les humains et les machines. Deuxièmement, les données sont considérées comme *accessibles* lorsque les utilisateurs peuvent les consulter à l'aide d'un protocole ouvert et universellement implémentable. Mais cela ne signifie pas que les données doivent être librement ouvertes sans restrictions. Parfois, les données peuvent être "FAIR" et non ouvertes. En d'autres termes, les données FAIR doivent être associées à des métadonnées qui spécifient les conditions dans lesquelles les données sont accessibles. Troisièmement, les données sont *interopérables* lorsque n'importe quel utilisateur peut facilement et d'une manière standardisée, les traiter sans avoir recours à une chaîne de traitement spécifique. Les sous-principes I peuvent être considérés comme les plus difficiles à réaliser et les plus importants pour être FAIR. Plus spécifiquement, ils indiquent que les données et les métadonnées doivent être représentées dans un langage formel, accessible, partagé et générique pour la représentation des connaissances. De plus, ces données doivent elles-mêmes utiliser des vocabulaires ou ontologies FAIR et inclure des références qualifiées à d'autres données et métadonnées. Il est clair que le Web sémantique et les technologies de données liées ont été identifiés parmi les meilleurs candidats à utiliser pour la représentation des connaissances, la lisibilité par machine et l'interopérabilité sur le Web, mais les principes FAIR ne peuvent pas être réduits au Web sémantique [3]. Enfin, les données sont *réutilisables* lorsqu'elles sont fournies avec des informations claires sur la licence et l'utilisation des données pour les humains et les machines. Elles doivent également être associées à des métadonnées et à une documentation riche qui détaillent leur provenance (spécifications des données, financement, cas d'utilisation, versions, processus expérimentaux, etc.).

Les principes FAIR sont décrits à un niveau générique et leur

mise en œuvre a volontairement été mise de côté au début. Cependant, avec le succès de la démarche, cette question d'implémentation des principes FAIR est devenue très importante, et elle doit se décliner au cas par cas pour chaque type d'objet digital. En Europe, par exemple, la mise en œuvre des principes FAIR est réalisée, en partie, par le programme European Open Science Cloud (<https://eosc-portal.eu>) de la Commission européenne, en particulier par le groupe d'experts sur les données FAIR. L'initiative GO FAIR ([www.go-fair.org](http://www.go-fair.org)) a également pour objectif de favoriser l'adoption des principes FAIR, notamment via la description et l'élaboration de profils de mise en œuvre ("FAIR Implementation Profile") et le déploiement de serveur d'accès aux données ("FAIR Data Points"). Parmi d'autres initiatives internationales qui encouragent l'adoption des principes FAIR, nous pouvons citer : le programme américain NIH Data Commons (<https://commonfund.nih.gov/commons>), la Research Data Alliance ([www.rd-alliance.org](http://www.rd-alliance.org)), plusieurs projets H2020 dont FAIRsFAIR ([www.fairsfair.eu](http://www.fairsfair.eu)), et des initiatives communautaires telles que le Food System GO FAIR Implementation Network ([www.foosin.fr](http://www.foosin.fr)).

Bien que la vision FAIR soit largement reconnue par plusieurs initiatives internationales et que les ontologies soient considérées comme un élément clé pour rendre les données FAIR, à ce jour il n'existe toujours pas de mécanisme, ni d'outil d'évaluation du niveau de "FAIRness" pour les ressources sémantiques. En effet, aucun des travaux existants dans la littérature n'a réussi à définir une approche de FAIRness claire et complète qui couvre à la fois les aspects méthodologiques et techniques associés aux ressources sémantiques. Pour toutes ces raisons, nous présentons ici le développement d'une méthodologie et d'un outil, très fortement inspiré de l'état de l'art actuel, pour l'évaluation du degré de mise en œuvre – nous parlerons aussi de niveau de "FAIRness" ou "FAIRness assessment" – des principes FAIR pour les ressources sémantiques. Notre approche se base sur une représentation riche et structurée des métadonnées d'une ontologie.<sup>1</sup> Nous nous basons sur le modèle de métadonnées MOD (voir Section 3) qui en amont a identifié les vocabulaires de métadonnées (et les propriétés définies dans ces vocabulaires) pour les ontologies. Nous avons développé un prototype de FAIRness assessment pour AgroPortal (<http://agroportal.lirmm.fr>) un portail pour les ressources sémantiques et les ontologies en agronomie [7].

L'article est structuré comme suit : la Section 2 présente l'état de l'art actuel sur l'évaluation du niveau de FAIRness. Ensuite, la Section 3 décrit les étapes de conception de notre méthodologie d'évaluation du niveau de FAIRness des ressources sémantiques basée sur les métadonnées, détaille notre approche en présentant une projection de chaque principe FAIR pour les ressources sémantiques, et propose des indicateurs de mesure présentés sous forme de questions. Puis, la Section 4 présente notre prototype implémenté dans AgroPortal, explique son fonctionnement, et fournit des résultats préliminaires pour quelques ressources sémantiques

dans le domaine de l'agriculture. Enfin, la Section 5 résume notre contribution et cite des perspectives.

## 2 Etat de l'art

Avant l'apparition des principes FAIR, en 2011, Berners-Lee a présenté les principes fondamentaux du Linked Open Data [8] pour rendre les données disponibles, partageables et interconnectées sur le Web. Les principes FAIR ont été proposés pour des raisons similaires en mettant davantage l'accent sur la réutilisabilité des données. Les principes LOD 5-star ont été spécialisés en 2014 pour les vocabulaires [9] et présentés sous la forme de cinq règles à suivre pour créer et publier des "bons" vocabulaires sur le Web. Dans ce schéma, les étoiles indiquent la qualité des données menant à une meilleure structure (e.g., l'utilisation des recommandations du W3C) et une meilleure interopérabilité pour la réutilisation (c'est-à-dire la représentation des métadonnées, la réutilisation des vocabulaires et l'alignement). Bien que le système de notation 5-star proposé pour les vocabulaires soit simple, à ce jour aucun outil de mise en œuvre n'a été développé autour de ces principes; ces principes ne sont pas non plus bien référencés dans la littérature. Une première étude d'alignement des principes LOD et FAIR a été réalisée dans [10], une deuxième étude plus approfondie a été proposée par Poveda et al.[11]; nous avons intégré ces alignements dans notre méthodologie.

En 2017, l'initiative *Minimum Information for Reporting an Ontology* (MIRO) a publié des directives destinées aux développeurs, pour la description d'une ontologie dans des rapports scientifiques [12]. Les directives MIRO visent à améliorer la qualité et la cohérence des descriptions du contenu de l'information; y compris la méthodologie de développement, la provenance et le contexte des informations de réutilisation. Ces directives définissent des éléments d'information (tels que nom, licence, URL) et spécifient leur niveau d'importance : 'must', 'should', 'optional'. Ce travail était significatif mais, jusqu'à présent, il n'existe aucune étude sur la façon dont les directives MIRO s'alignent avec ou complètent les principes FAIR. Le modèle de métadonnées MOD 1.4<sup>2</sup> a cependant fourni un alignement entre chaque directive MIRO et les propriétés de métadonnées correspondantes dans MOD. Nous avons donc utilisé cet alignement dans notre méthodologie, pour influencer le score de FAIRness avec les directives MIRO. Par exemple, la 'guideline' A.3 de MIRO recommande de préciser la licence de l'ontologie, et MOD suggère d'utiliser `dct:license` pour cela. Ici, nous ajoutons que cela implémente le principe R1.1.

Depuis 2018, plusieurs approches et outils génériques pour l'auto-évaluation du niveau de FAIRness d'une ressource sont apparues; nous pouvons citer : SHARC[13] , FDMM [14], FAIR Metrics [15], [16] (devenu FAIR evaluator), FAIRdat [17] et FAIR-Aware [18]. Une analyse synthétique de comparaison de 12 outils d'évaluation faite par la Research Data Alliance est disponible sur ce lien.<sup>3</sup> Par manque d'espace, nous ne présentons que l'approche FDMM et SHARC; les

<sup>1</sup> Dans cet article, nous utilisons parfois le terme "ontologie" pour faire référence à des ressources sémantiques [4], des systèmes d'organisation des connaissances [5], ou des artefacts sémantiques [6]. Typiquement, des terminologies, thésaurus et vocabulaires.

<sup>2</sup> Metadata for Ontology Description and Publication Ontology: <https://github.com/sifproject/MOD-Ontology>

<sup>3</sup> <https://github.com/rd-alliance/FAIR-data-maturity-model-WG>

approches génériques intégrées dans notre méthodologie.

En 2017, le groupe d'intérêt SHARC (SHARing Rewards and Credit) de la Research Data Alliance (RDA) a proposé une grille d'évaluation du niveau de FAIRness dédiée aux données. L'objectif de ce travail est de guider les chercheurs et les autres parties prenantes à adopter les principes FAIR. La grille FAIR définit un ensemble de 45 critères génériques avec trois niveaux d'importance ('essential', 'recommended', 'desirable'); les critères sont formulés en questions qui sont parfois dépendantes les unes des autres comme dans un arbre de décision.

En 2018, le groupe de travail RDA FAIR Data Maturity Model (FDMM) a publié une liste de recommandations génériques visant à normaliser la méthodologie d'évaluation du niveau de FAIRness afin de permettre la comparabilité des résultats. Plus concrètement, le travail est présenté sous forme d'une grille qui définit 47 critères génériques avec des priorités ('essential', 'important', 'useful'); ces critères sont dérivés de chaque principe FAIR.

En résumé, la grille SHARC et FDMM considèrent que certains principes FAIR sont plus importants que d'autres. Nous avons suivi cette vision dans notre méthodologie et avons décidé d'inclure les résultats de SHARC et FDMM pour nous aider à déterminer l'ordre d'importance des principes FAIR pour les ressources digitales.

Ces deux dernières années, des approches plus spécifiques autour de l'évaluation des ressources sémantiques selon les principes FAIR, ont été publiées, nous les détaillons ci-dessous.

En mars 2020, le projet H2020 FAIRsFAIR a introduit 17 recommandations pour le partage des ressources sémantiques avec les principes FAIR [6]. Pour chaque recommandation, les auteurs fournissent une description détaillée et soulignent les technologies du Web sémantique qui peuvent être utilisées pour la mise en œuvre de cette recommandation pour n'importe quelle ressource sémantique. Ce travail, très pertinent pour nous, et en cours de révision par la communauté scientifique, reste préliminaire, et devra être complété par des indicateurs concrets de FAIRness qui pourront être utilisés pour évaluer automatiquement le degré de mise en œuvre des recommandations proposées pour les ressources sémantiques.

Également, Garijo et Poveda [11] proposent une liste de bonnes pratiques pour des ontologies FAIR. Dans un autre article [19], les auteurs complètent leur travail par une analyse qualitative du degré d'alignement des principales approches existantes sur le partage des ressources sémantiques (5-stars LOD [8], 5-stars V [9], et FAIRsFAIR) avec les principes FAIR. Cette analyse démontre qu'aucune approche existante n'est complètement alignée aux principes FAIR et soulève aussi des enjeux scientifiques importants. L'approche de Poveda et al. est significative mais elle est limitée à l'étude de certaines approches sémantiques et ne propose pas des métriques de mesure du niveau de FAIRness. Nous pensons que ce travail d'analyse réalisée est important mais il devrait être amélioré pour inclure d'autres approches plus génériques notamment FDMM et SHARC. C'est également une contribution de notre travail présenté ici.

Finalement, l'outil DBPedia Archivio lancé par Frey et al. [20] est le premier outil public pour les ontologies FAIR. Archivio

propose l'archivage et l'évaluation du niveau de FAIRness de toute ontologie hébergée dans sa propre librairie publique. L'inconvénient de cet outil est qu'il ne détaille pas ses résultats d'évaluation et ne propose pas de recommandations pour guider ses utilisateurs dans l'amélioration du score de FAIRness obtenu pour leurs ontologies. D'un point de vue méthodologique, Archivio n'a pas encore dévoilé l'approche qu'il applique dans le processus d'évaluation de ses ressources. Par conséquent, leur approche reste inconnue pour la communauté Web sémantique.

Notre analyse de l'état de l'art montre clairement qu'aucune *approche spécifique* pour les ontologies n'est strictement alignée aux 15 sous-principes FAIR (voir Tableau 1); également qu'aucune *approche générique* des principes FAIR n'est strictement applicable aux ontologies. Ainsi, nous proposons ici une méthodologie qui fait converger les deux visions. Nous constatons aussi que les approches génériques tiennent fortement compte de tous les sous-principes FAIR sauf F2, F3 et A2. Cependant, les approches spécifiques accordent plus d'importance à F1, A1, I1, I2, R1 et R.1.2. Les résultats obtenus soulignent aussi la nécessité d'établir un ensemble minimum de représentation des métadonnées de l'ontologie (en cours de discussion dans le cadre du projet FAIRsFAIR), une meilleure fédération et interopération des portails et services pour les ontologies, des stratégies de maintenance long terme pour les ontologies et leurs métadonnées au sein de ces portails/services, et des meilleures pratiques pour documenter et communiquer sur les ontologies.

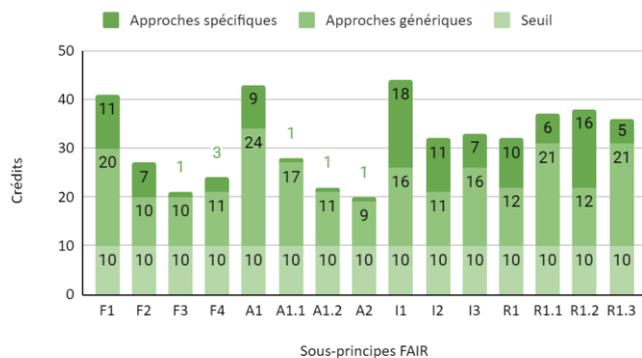
Pour construire notre approche, nous avons considéré SHARC, FDMM, LOD 5-star, MIRO, FAIRsFAIR et Poveda et al. Nous avons étudié chaque approche (i) en analysant comment elle s'aligne avec chaque sous-principe FAIR (voir Tableau 1) et (ii) en pondérant (à l'aide de crédits) les sous-principes en fonction de l'importance que lui a donné une approche. Plus précisément, nous avons fixé une valeur "de départ" par défaut à 10 crédits égale pour chaque sous-principe. Dans notre calcul, nous sommes partis de cette valeur de départ et avons ajouté des crédits aux sous-principes en fonction de l'importance que leur donnent les approches étudiées. Le calcul des crédits obtenus pour chaque sous-principe n'est pas détaillé ici mais peut être consulté dans [21]. La Figure 1 montre les valeurs de crédits obtenus pour chaque sous-principe. Le vert clair est la valeur de départ; le vert moyen représente les ajouts de crédits obtenus avec les approches génériques, et le vert foncé représente les ajouts de crédits obtenus avec les approches spécifiques aux ontologies. Ainsi, on peut constater que les trois principes qui ressortent le plus sont F1, A1 et I1 et ceux qui ressortent le moins sont F3, A1.2 et A2.

	SHARC	FDMM	MIRO	5*V	FsF	Poveda et al.
<b>F1</b>	X	X	X	-	X	X
<b>F2</b>	X	X	X	-	X	X
<b>F3</b>	X	X	-	-	X	-
<b>F4</b>	X	X	-	-	X	X
<b>A1</b>	X	X	X	X	X	X
<b>A1.1</b>	X	X	-	-	X	-

A1.2	X	X	-	-	X	-
A2	X	X	-	-	X	-
I1	X	X	X	X	X	X
I2	X	X	X	-	X	-
I3	X	X	X	X	X	X
R1	X	X	X	X	-	X
R1.1	X	X	X	-	X	X
R1.2	X	X	X	-	X	X
R1.3	X	X	-	-	X	X

**Tableau 1.** Aligement des approches existantes aux 15 sous-principes FAIR. Le symbole (X) indique que l'approche étudie le principe concerné, en revanche le symbole (-) indique le contraire.

Distribution des crédits par sous-principe FAIR



**Figure 1.** Distribution des crédits obtenus pour chaque sous-principe FAIR; le calcul est issu de l'intégration des approches spécifiques, génériques, et du seuil.

### 3 Méthodologie

Dans ce travail, nous considérons que l'évaluation du niveau de FAIRness des ontologies et des ressources sémantiques devrait autant que possible être basée sur l'évaluation de descriptions formelles de leurs métadonnées, idéalement indexées, partagées et normalisées par des portails d'ontologies tels que le NCBO BioPortal [22] ou Ontology Lookup Service de EBI [23] en biomédecine, ou AgroPortal en agronomie. Dans des travaux antérieurs, nous avons construit un nouveau modèle de métadonnées harmonisé pour AgroPortal et démontré qu'il améliore l'identification et la réutilisation des ressources sémantiques en agronomie [7], [24]. Ici, nous démontrons que les portails d'ontologies sont également importants pour l'évaluation du niveau de FAIRness.

Dans cette section, nous interprétons trois sous-principes (F1, A1, I1 et R1.1) pour les ressources sémantiques et listons les propriétés de métadonnées standardisées utilisées pour décrire les informations nécessaires liées à l'évaluation de ces sous-principes. Nous traitons dans cet article trois sous principes sur 15, le reste de la méthodologie est documentée sur GitHub, dans le dépôt du code open source qui implémente notre approche: <https://github.com/agroportal/fairness>. Nous avons utilisé le modèle MOD comme référence, il regroupe 346 propriétés

tirées de 23 vocabulaires de métadonnées (tels que Dublin Core, DCAT, VoID, ADMS, VOF, Schema.org, etc.) [24]. MOD n'est donc pas un standard, mais un listing de standards. Ici, nous utilisons le modèle de métadonnées MOD pour identifier sans ambiguïté quelle propriété peut être utilisée; cependant, notre méthodologie est indépendante de ce modèle et exige seulement que l'information soit bien représentée et harmonisée dans les métadonnées d'une ontologie.

Nous avons distribué les crédits obtenus précédemment sur des séries de questions proposées pour l'évaluation de chaque sous-principe FAIR. La distribution des crédits d'un sous-principe sur les questions associées est basée sur une distribution équilibrée et pertinente qui pourrait être modifiée ou ajustée si des questions étaient supprimées ou ajoutées à l'avenir. Toutes nos questions d'évaluation, sauf celles du sous-principe R1.3, sont génériques (c'est-à-dire non spécifiques à l'agronomie ou à AgroPortal) et pourront être appliquées sur n'importe quelle ontologie par d'autres communautés scientifiques. Nous recommandons donc aux futurs utilisateurs de notre travail de spécifier eux même les questions de R1.3 en s'adressant aux experts de leur domaine d'intérêt.

Concrètement, le résultat de notre travail est une grille d'évaluation composée de 68 questions couvrant la totalité des sous-principes FAIR; nous avons identifié 55 propriétés de métadonnées nécessaires (correspondantes à 309 crédits) pour répondre à toutes ces questions sauf celles de F2; 73 propriétés additionnelles sont proposées dans MOD pour enrichir une ontologie avec des métadonnées et augmenter son niveau de FAIRness. En d'autres termes, nous proposons dans F2 d'évaluer toutes les propriétés de métadonnées qui ne sont pas considérées dans le reste des sous-principes. Pour chaque sous-principe, nous avons attribué à chaque question un certain nombre de crédits dans la limite des valeurs identifiées dans la Figure 1. Il faut noter que les questions sont majoritairement binaires (oui ou non) ainsi, pour une question donnée une ressource sémantique se verra en général attribuer soit tous les crédits, soit aucun; mais rarement une valeur entre les deux.

Dans la suite, nous listons les questions d'évaluation de F1, A1, I1 et R1.1 (c'est-à-dire, une illustration d'un sous-principe pour un groupe de principe donné) et précisons le nombre de crédits que notre algorithme attribuera si le test de conformité sémantique est validé pour chaque sous-principe. Le score final de FAIRness est ensuite une simple somme de l'ensemble des crédits obtenus aux questions d'évaluation. Ce score peut être normalisé sur 100 pour faciliter sa compréhension et la comparaison des ontologies entre elles.

**Sous-principe F1. Les ontologies et les métadonnées d'ontologie reçoivent un identifiant global unique et persistant.** F1 concerne les *identifiants*; bien que cela n'est pas explicitement mentionné dans l'intitulé du sous-principe, plusieurs travaux demandent que les identifiants soient résolubles, c'est-à-dire un identifiant qui permettent d'avoir accès à la ressource ou à la description de la ressource (déréférencement). FAIRsFAIR recommande l'utilisation d'un identifiant [25] *globalement unique, persistant et résolvable* (GUPRI) pour les ontologies et les métadonnées d'ontologies.

Dans la plupart des cas, les ontologies décrites avec les langages du Web sémantique possèdent un *identifiant de ressource unique* (URI) et les métadonnées sont soit représentées au

niveau du fichier de l'ontologie (le plus fréquent) soit dans un fichier externe. Les URIs sont généralement uniques sur le Web, mais ils ne sont pas toujours pérennes et résolubles. Dans certains cas, les ontologies peuvent avoir un identifiant supplémentaire attribué par une organisation externe telle qu'un DOI. Parfois, les URIs prennent la forme de PURL qui sont supposés persistants mais qui ne sont pas certifiés comme le sont les DOI. Idéalement, les URIs d'ontologies devraient pouvoir être résolubles pour garantir un degré de conformité plus élevé pour F1. Lorsqu'elles sont sauvegardées dans un fichier séparé, les mêmes règles concernant l'identification doivent s'appliquer au fichier de métadonnées de l'ontologie. En complément, les communautés qui développent des ontologies essaient parfois de maintenir une utilisation cohérente des acronymes pour identifier les ontologies. Par exemple, dans le cadre de l'OBO Foundry, un nom court est obligatoire demandé et sera utilisé pour identifier l'ontologie et pour l'attribution du PURL (par exemple, l'acronyme de l'*Agronomy Ontology* est AGRO et son PURL est <http://purl.obolibrary.org/obo/agro.owl>)

F1 peut être évalué en vérifiant la valeur affectée à la propriété `owl:ontologyIRI`, utilisée pour coder l'URI de l'ontologie et la propriété `dct:identifier`, utilisée pour coder un autre identifiant "externe". De plus, la propriété `owl:versionIRI`, qui sauvegarde un URI spécifique à la version, peut également être utilisée pour évaluer si l'ontologie distingue clairement des identifiants de versions. La Liste 1 résume les questions d'évaluation pour ce sous-principe.

#### Liste 1. Questions d'évaluation de F1 (41 crédits)

- Q1. Une ontologie a-t-elle un identifiant "local", c'est-à-dire un identifiant globalement unique et potentiellement persistant mais attribué par le développeur (ou l'organisation de qui développe l'ontologie) ? **9 cts**
- Q2. Une ontologie fournit-elle un identifiant "externe" supplémentaire, c'est-à-dire un identifiant globalement unique et persistant attribué par un organisme accrédité ? **6 cts**
- Q3. Si oui, cet identifiant externe est-il un DOI ? **5 cts**
- Q4. Les métadonnées de l'ontologie sont-elles incluses dans le fichier d'ontologie et partagent-elles par conséquent les mêmes identifiants ? **6 cts**
- Q5. Sinon, l'ensemble des métadonnées est-il clairement identifié par son propre GUPRI ? **6 cts**
- Q6. Une ontologie fournit-elle un URI spécifique à la version ? **4 cts**
- Q7. Si oui, cet URI est-il résolvable/ déréférencable ? **5 cts**

**Sous-principe A1. Les ontologies et les métadonnées d'ontologie sont accessibles par leur identifiant à l'aide d'un protocole de communication normalisé.** A1 exprime l'importance des identifiants pour rendre une ressource accessible sur le Web. En Web sémantique, une ontologie hébergée sur un serveur Web devient accessible via le protocole standard HTTP. Ainsi, les objets d'une ontologie, (c'est-à-dire, classes, relations et métadonnées) peuvent être facilement récupérées via des services Web et dans le meilleur des cas elles peuvent aussi être disponibles sous différents formats (par exemples, JSON, HTML, texte, etc.) en utilisant la négociation de contenu.

A1 peut-être évalué en vérifiant la résolubilité (et la prise en charge de la négociation du contenu) via HTTP des URIs d'ontologie et des métadonnées d'ontologie. Voir les métadonnées F1 liées aux identifiants pour l'évaluation de l'accessibilité sous HTTP. Dans le schéma MOD, d'autres propriétés existent pour évaluer si une ontologie et les métadonnées d'ontologie sont accessibles à travers d'autres protocoles de communication comme les requêtes SPARQL : la propriété `sd:endpoint` peut être utilisée pour stocker le point d'accès SPARQL qui peut être utilisé pour récupérer le contenu et les métadonnées de l'ontologie. De plus, les propriétés `void:openSearchDescription` et `void:uriLookupEndpoint` peuvent également être utilisées pour évaluer l'existence d'un moteur de recherche en text libre ou par URI sur l'ontologie; cependant nous définissons ces propriétés comme des métadonnées optionnelles dans F2.

#### Liste 2. Questions d'évaluation de A1 (43 crédits)

- Q1. Est-ce que l'URI (ou d'autres identifiants) de l'ontologie se déréférence vers l'ontologie ? **6 cts**
- Q2. Est-ce que l'URI de l'ontologie ou l'URI de métadonnées externes se déréférence vers les valeurs de métadonnées ? **7 cts**
- Q3. L'ontologie et ses métadonnées prennent-elles en charge la négociation de contenu ? **24 cts**
- Q4. Une ontologie et ses métadonnées sont-elles accessibles via un autre protocole standard tel que SPARQL ? **6 cts**

**Sous-principe II. Les ontologies et les métadonnées d'ontologie utilisent un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances.** Il met l'accent sur l'importance du langage de représentation des connaissances. Une ontologie est généralement une ressource conçue pour être compréhensible par la machine et qui repose donc sur un *langage formel*. Cependant, certaines ontologies ou ressources sémantiques peuvent être décrites sous forme textuelle ou graphique qui ne sont pas directement exploitables par une machine.

Théoriquement, une ontologie ou une ressource sémantique est sauvegardée dans un fichier en utilisant une *syntaxe* dédiée (RDF/XML, Turtle, JSON-LD) et un *langage de représentation* (OWL, SKOS, RDFS, OBO). Une ressource sémantique peut avoir différents niveaux de formalité (par exemples, ontologie, terminologie, thésaurus, vocabulaire). Les métadonnées d'ontologies sont généralement représentées en utilisant la même syntaxe et le même langage de représentation que l'ontologie elle-même. Lorsque les métadonnées sont sauvegardées dans un fichier externe, le langage de représentation des connaissances doit être évalué indépendamment.

Le sous-principe II peut être évalué en regardant le niveau de formalité et d'accessibilité du langage de représentation utilisé pour encoder l'ontologie ainsi que dans quelle mesure ce langage est partagé/adopté par une communauté, s'il est générique (c'est-à-dire "largement applicable" et non spécifique au domaine) et s'il est recommandé par des organismes de normalisation (dans notre cas le W3C principalement). Dans notre méthodologie, nous avons considéré que les ressources sémantiques et leurs métadonnées

peuvent être décrites en utilisant le langage OWL, OBO, RDFS et SKOS (indépendamment de leur syntaxe associée) ainsi que dans les formats CSV, XML, PDF ou TEXT. Pour l'évaluation de I1, nous pouvons regarder le langage de représentation, et le niveau de formalité utilisé et enregistré en tant que métadonnée (déclaré par les auteurs ou le portail d'hébergement). Nous pouvons aussi regarder la disponibilité de la ressource dans d'autres formats/syntaxes. Dans le schéma MOD, le langage de représentation d'une ontologie est représenté via la propriété `omv:hasOntologyLanguage`, son niveau de formalité est décrit avec la propriété `omv:hasFormalityLevel`, sa syntaxe est décrite avec la propriété `omv:hasOntologySyntax`. Si une ressource sémantique est disponible dans un autre format ou une autre syntaxe, ces informations peuvent être décrites avec les propriétés `dct:hasFormat` et `dct:isFormatOf`. La Liste 3 résume les questions d'évaluation pour ce sous-principe.

### Liste 3. Questions d'évaluation de I1 (44 crédits)

- Q1.** Quel est le langage de représentation utilisé pour l'ontologie et les métadonnées de l'ontologie ? **20 cts\***
- Q2.** Le langage de représentation utilisé est-il une recommandation du W3C ? **10 cts**
- Q3.** La syntaxe de l'ontologie est-elle déclarée ? **5 cts**
- Q4.** Le niveau de formalité de l'ontologie est-il déclaré ? **5 cts**
- Q5.** La disponibilité d'autres formats est-elle déclarée ? **4 cts**
- (\*) Nous proposons l'échelle suivante pour la notation de chaque langage de représentation : (owl, 20 pts) - (skos, 18 pts) - (rdfs, 16 pts) - (obo, 14 pts) - (xml, 12 pts) - (csv, 11 pts) - (pdf, 5 pts) - (txt, 5 pts). Nous donnons une petite avance à OWL qui d'après nous est le langage le plus "formal and broadly applicable".

**Sous-principe R1.1. Les ontologies et les métadonnées d'ontologie sont publiées avec une licence d'utilisation claire et accessible.** Bien que l'ouverture ne soit pas un critère obligatoire pour rendre les données FAIR, il est évident que le fait de rendre les ontologies et les métadonnées d'ontologies ouvertement et librement disponibles va améliorer leur réutilisation. Quel que soit le type de licence choisie, R1.1 exige une représentation compréhensible par la machine pour la licence. En fait, l'absence d'une description explicite de la licence pourrait empêcher des personnes de réutiliser l'ontologie, même si elle était à l'origine ouverte et destinée à être partagée. Actuellement, le site RDF License<sup>4</sup> offre des URIs et des descriptions RDF pour la plupart des licences ; également le vocabulaire Creative Commons<sup>5</sup> fournit plusieurs propriétés pour garantir une description compréhensible par la

machine des droits d'accès et des licences.

Le sous-principe R1.1 peut être évalué en vérifiant si les informations de licence et de droits d'accès sont fournies et résolubles (notamment la licence). Le modèle de métadonnées MOD suggère les propriétés `dct:license` pour décrire les informations de licence, et `dct:accessRights` pour détailler les droits d'accès (qui a accès à quoi). Le modèle MOD propose aussi des propriétés pour décrire les informations sur les permissions et les conditions d'utilisation associées à l'ontologie (`cc:morePermissions`, `cc:useGuidelines`) ainsi que le détenteur du droit d'auteur (`dct:rightsHolder`). Dans notre méthodologie, nous supposons que l'ontologie et ses métadonnées sont régies par la même licence – par défaut lorsque les métadonnées sont effectivement décrites dans le même fichier que l'ontologie – mais bien sûr, si ce n'est pas le cas, deux licences doivent être spécifiées et les crédits doivent être divisés en fonction. La Liste 4 résume les questions d'évaluation pour ce sous-principe.

### Liste 4. Questions d'évaluation de R1.1 (37 crédits)

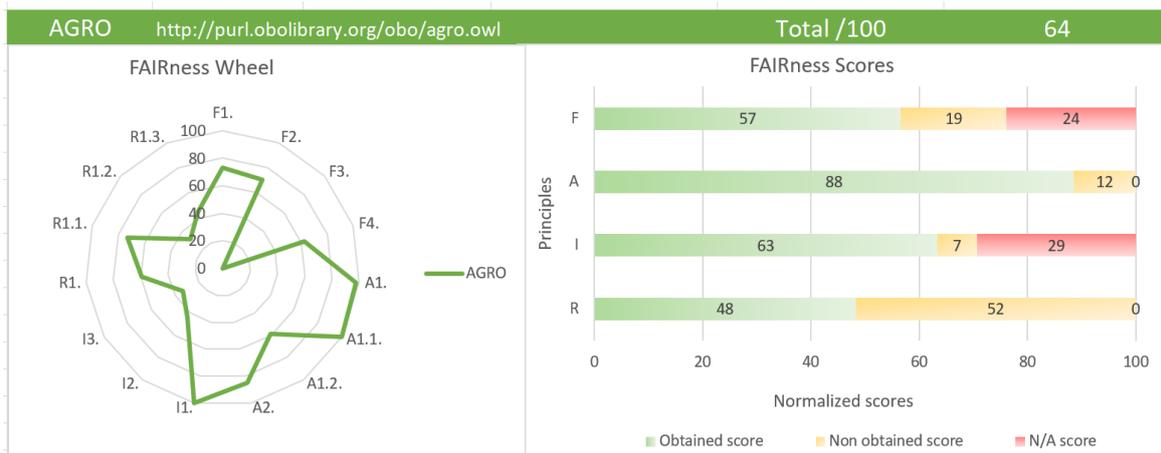
- Q1.** La licence de l'ontologie est-elle clairement spécifiée (c'est-à-dire avec un identifiant unique et persistant) ? **8 cts**
- Q2.** Si oui, la description de la licence est-elle accessible et résolvable par une machine ? **7 cts**
- Q3.** Les droits d'accès à l'ontologie sont-ils clairement spécifiés / déclarés ? **7 cts**
- Q4.** Les autorisations, les conditions d'utilisation et le détenteur des droits d'auteur sont-ils clairement documentés ? **15 cts**

## 4 Résultats

Nous avons implémenté un prototype d'évaluation du niveau de FAIRness sous la forme d'un service Web indépendant qui utilise – via l'API REST – les métadonnées des ressources sémantiques dans AgroPortal et évalue automatiquement 59 questions sur les 68 définies dans notre approche. Seulement 10 questions ne peuvent pas encore être évaluées dans AgroPortal : 1 question liée à F1, 3 questions liées à F3, 3 questions I2, et 3 questions I3. L'extension du modèle MOD avec des nouvelles propriétés devrait aider à couvrir trois de ces questions e.g., évaluer le lien entre l'ontologie et les métadonnées (F1-Q5), décrire l'état de curation (I2-Q6) ou la qualification des alignements (I3). AgroPortal étant un portail d'hébergement des ressources sémantiques donc ici, il est important de noter que notre service Web traite les métadonnées de l'objet stocké chez AgroPortal et non pas celles du fichier d'origine.

<sup>4</sup> <http://rdflicense.appspot.com/>

<sup>5</sup> <https://creativecommons.org/licenses/>



**Figure 2.** Synthèse graphique obtenue pour l'évaluation du niveau de FAIRness de l'Agronomy Ontology (AGRO).

Le service Web prend comme paramètre un acronyme<sup>6</sup> d'une ressource sémantique, c'est-à-dire un identifiant local dans AgroPortal, et renvoie en sortie un fichier JSON qui contient le score FAIR obtenu pour chaque question et chaque sous-principe (tableau 'scores'), ainsi que les scores par principe (champ numérique 'total score'). Chaque score de question est justifié par une petite phrase de justification (tableau 'explanations'), afin que l'utilisateur puisse être informé de la raison pour laquelle ce score a été obtenu. Ci-après, un exemple d'appel au service Web pour l'évaluation de l'ontologie AGRO (Agronomy Ontology) :

[http://services.agroportal.lirmm.fr/fairness\\_assessment/?portal=agroportal&ontology=AGRO](http://services.agroportal.lirmm.fr/fairness_assessment/?portal=agroportal&ontology=AGRO)

Le temps de traitement du service Web est non linéaire, l'évaluation des questions ne dépend pas de la taille de l'ontologie (classes/relations) ; elle repose uniquement sur une liste déterminée de métadonnées décrivant l'ontologie. Un exemple de résultat pour le critère F1 est illustré dans la Figure 3.

A court terme, nous prévoyons d'enrichir notre sortie JSON avec des scores normalisés sur 100 (tel que rapportés sur la Figure 2) et de permettre la visualisation des résultats de FAIRness sur la page 'Summary' d'AgroPortal : (e.g., <http://agroportal.lirmm.fr/ontologies/AGRO>). Sur du plus long terme, il s'agira de représenter nos résultats dans un format structuré standard pour la représentation de FAIRness assessment; de tels formats sont actuellement en cours de discussion.

Également, le service Web prend comme paramètre le portail à utiliser pour l'évaluation du niveau de FAIRness (paramètre `portal`) car notre objectif est d'offrir un service générique pour tout portail d'ontologies mettant en œuvre le modèle de métadonnées d'ontologie MOD et/ou offrant un modèle de métadonnées riche et harmonisé pour les ontologies. Le code du prototype actuel est basé sur la version 3 d'OntoPortal (<https://ontportal.org>) dans lequel nous avons changé le modèle de métadonnées [24].<sup>7</sup> Ce code est utilisé pour AgroPortal et le SIFR BioPortal (un portail d'ontologies et de

terminologies biomédicale françaises) [26]; ainsi, pour l'instant, seul AgroPortal et le SIFR BioPortal supporte le service Web de FAIRness assessment implémenté. Dans le cadre de l'Alliance OntoPortal, nous envisageons d'étendre notre modèle de métadonnées aux autres portails d'ontologies basés sur cette technologie (i.e., NCBO BioPortal, LifeWatch EcoPortal, MedPortal et MatPortal).

```
{
  "AGRO": {
    "Findable": {
      "F1": {
        "resultSet": {
          "explanations": [
            "Valid ontology URI",
            "Resolvable ontology URI",
            "Valid GUID",
            "GUID is not a DOI",
            "Metadata are not included in the ontology file",
            "Metadata are identified by a resolvable URI",
            "Valid URI version",
            "Resolvable ontology URI version"
          ],
          "scores": [
            3,
            6,
            6,
            0,
            0,
            6,
            4,
            5
          ]
        },
        "totalScore": 30
      }
    }
  }
}
```

**Figure 3.** Résultat d'évaluation de F1 pour l'ontologie AGRO. La sortie JSON montre (a) les explications données aux scores, (b) les détails des scores et (c) le score total du sous-principe concerné.

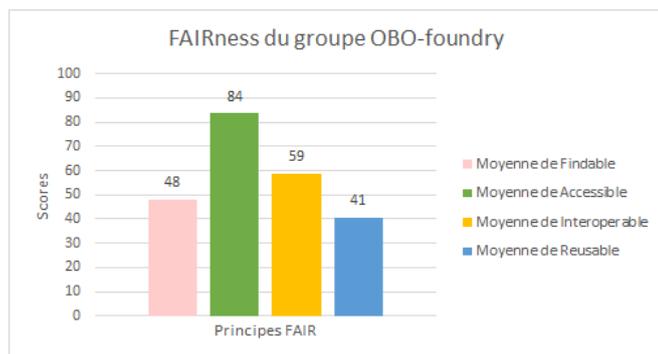
Pour analyser les scores de FAIRness par ontologie ou groupe

<sup>6</sup> La liste des acronymes des ontologies hébergées dans AgroPortal est disponible via le lien : <http://agroportal.lirmm.fr/ontologies/>

<sup>7</sup> <https://github.com/ontportal-lirmm>

d'ontologies, nous avons produit dans une feuille Excel des graphiques synthétiques pour visualiser les résultats. La Figure 2 détaille à titre d'exemple le score obtenu pour l'ontologie AGRO. La "roue de FAIRnes" de la Figure 2 indique la répartition du score obtenu par l'ontologie ou par un groupe d'ontologies sur l'ensemble des 15 sous-principes. Le score global normalisé pour AGRO est de 64/100). L'histogramme de la Figure 2 affiche pour chaque principe : le score total obtenu (série en vert), les crédits non obtenus dans AgroPortal (série jaune), et les crédits qui ne peuvent pas encore être affectés/calculés au sein de AgroPortal (série en rouge). A titre d'exemple, l'ontologie AGRO a un score normalisé de 57 sur les 76 points évaluables sur AgroPortal pour le principe F.

Les moyennes que nous avons pu obtenir sur l'ensemble des ontologies d'AgroPortal nous indique qu'un score au-dessus de 65 est relativement un "bon" score de FAIRness. En effet, la moyenne des scores pour les 134 ontologies d'AgroPortal que nous avons testées en mars 2021 est de 50. La médiane est de 49. Etant donné qu'aucun des travaux existants ne définissent de valeur de référence pour définir les niveaux de FAIRness, nous nous reposons ici sur nos statistiques expérimentales comme la valeur moyenne de FAIRness. Nous expliquons dans [21] qu'une métrique est indispensable pour justement indiquer à partir de quand une ontologie n'est pas FAIR, est FAIR ou même FAIRer.



**Figure 4.** Synthèse de l'évaluation du degré de mise en œuvre des principes FAIR au sein du groupe 'OBO-foundry'.

La Figure 4 montre le résultat d'évaluation du degré de mise en œuvre des principes FAIR pour 24 ontologies appartenant au groupe "OBO-foundry" dans AgroPortal. Le score de FAIRness pour tout le groupe est de 58 (c'est la moyenne des scores décrits en Figure 4) : l'ontologie qui obtient le meilleur score de FAIRness est la *Phenotype And Trait Ontology* (PATO) avec un score de 65 et un des scores le moins haut est obtenu pour l' *Agriculture and Forestry Ontology* (AFO). La valeur de FAIRness du groupe OBO est plutôt satisfaisante; cette valeur, au dessus de la moyenne est essentiellement obtenue grâce: (i) aux principes de conception que l'OBO Foundry demande aux ontologies qu'elle héberge (incluant des aspects sur l'utilisation des PURLs, la maintenance des fichiers de métadonnées, la clarté sur les conditions d'utilisation, et le support de négociation du contenu sous différents formats); et (ii) à l'hébergement de ces ontologies dans un 'repository' comme AgroPortal qui aide à implémenter certains principes FAIR pour n'importe laquelle des ontologies hébergées (e.g., accessibilité du contenu avec le protocole HTTP, la description riche des métadonnées, les fonctionnalités de recherche,

l'alignement avec d'autres ressources, et l'archivage des versions, etc.). Néanmoins, ces ontologies peuvent devenir encore plus FAIR en améliorant la mise en œuvre des sous-principes tels que F2 et R1.2 à travers les propriétés de métadonnées MOD relatives à ces sous-principes.

Une analyse détaillée du niveau de FAIRness de toutes les ontologies de AgroPortal fera l'objet d'une autre communication.

## 5 Discussion et conclusions

Ce travail aborde la problématique d'évaluation de la mise en œuvre des 15 principes FAIR pour les ressources sémantiques et apporte des solutions concrètes qui faciliteraient l'adoption de ces principes par la communauté sémantique. Nous avons présenté une méthodologie et un outil d'évaluation du niveau de FAIRness. La méthodologie proposée est : (i) alignée aux approches de l'état-de-l'art, (ii) basée sur un modèle de métadonnées, (iii) générique et peut donc être appliquée pour tout type de ressource sémantique quel que soit le domaine d'application. L'implémentation que nous avons produite peut être subjective sur certains aspects (e.g., nombre de crédits par questions, liste des questions), c'est pourquoi nous avons voulu la méthodologie sous-jacente aussi générique que possible de façon à ce que chacun puisse la déployer avec ses spécificités. A termes, d'autres indicateurs (vote, usage, sondage) interviendront pour "évaluer" les outils de FAIRness assessment. Les résultats préliminaires de notre prototype montrent l'intérêt d'une évaluation par crédits pour l'ensemble des sous-principes et l'étude donne également une idée sur les analyses par ontologie ou groupe d'ontologies que nous pourrions générer à partir des scores obtenus.

Plusieurs enjeux scientifiques liés à l'évaluation du niveau de FAIRness des ressources sémantiques nécessitent d'être traités par la communauté nous citons à titre d'exemples : le besoin d'établir un consensus pour garantir la persistance des URIs (par exemple, un service d'enregistrement d'identifiants), de proposer un ensemble de métadonnées à évaluer pour chaque sous-principe, de fournir des mécanismes de standardisation et d'échange des métadonnées afin de faciliter leur récupération par les moteurs de recherche.

Prochainement, (i) nous effectuerons une analyse détaillée des scores de FAIRness pour l'ensemble des ontologies d'AgroPortal, (ii) nous réaliserons une enquête pour déterminer comment notre approche a aidé nos utilisateurs dans la sélection et l'amélioration des ressources sémantiques, (iii) nous continuerons nos efforts pour la standardisation et l'interopérabilité des métadonnées des ressources sémantiques au sein des initiatives internationales de la communauté FAIR (RDA, GO FAIR et projet FAIRsFAIR).

## Remerciements

Ce travail a été réalisé dans le cadre du projet ANR *Des Données aux Connaissances en Agronomie et Biodiversité* (D2KAB – [www.d2kab.org](http://www.d2kab.org) – ANR-18-CE23-0017) et du projet ANR *Participation française au GO FAIR Food Systems Implementation Network* (FooSIN – [www.foosin.fr](http://www.foosin.fr) – ANR 19-DATA-0019). Nous remercions également le groupe de travail VSSIG (*Vocabulary and Semantic Services Interest Group*) de

la Research Data Alliance ainsi que le projet H2020 FAIRsFAIR pour les discussions sur les ontologies et les vocabulaires FAIR.

## Références

- [1] M. D. Wilkinson *et al.*, « The FAIR Guiding Principles for scientific data management and stewardship », *Sci. Data*, vol. 3, n° 1, Art. n° 1, mars 2016, doi: 10.1038/sdata.2016.18.
- [2] B. Mons, *Data Stewardship for Open Science: Implementing FAIR Principles*. CRC Press, 2018.
- [3] B. Mons, C. Neylon, J. Velterop, M. Dumontier, L. O. B. da Silva Santos, et M. D. Wilkinson, « Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud », *Inf. Serv. Use*, vol. 37, n° 1, p. 49-56, janv. 2017, doi: 10.3233/ISU-170824.
- [4] C. Caracciolo *et al.*, « 39 Hints to Facilitate the Use of Semantics for Data on Agriculture and Nutrition », *Data Sci. J.*, vol. 19, n° 1, Art. n° 1, déc. 2020, doi: 10.5334/dsj-2020-047.
- [5] M. L. Zeng et P. Mayr, « Knowledge Organization Systems (KOS) in the Semantic Web: a multi-dimensional review », *Int. J. Digit. Libr.*, vol. 20, n° 3, p. 209-230, sept. 2019, doi: 10.1007/s00799-018-0241-2.
- [6] Y. Le Franc, J. Parland-von Essen, L. Bonino, H. Lehväsäläho, G. Coen, et C. Staiger, « D2.2 FAIR Semantics: First recommendations », mars 2020, doi: 10.5281/zenodo.3707985.
- [7] C. Jonquet *et al.*, « AgroPortal: A vocabulary and ontology repository for agronomy », *Comput. Electron. Agric.*, vol. 144, p. 126-143, janv. 2018, doi: 10.1016/j.compag.2017.10.012.
- [8] C. Bizer, T. Heath, et T. Berners-Lee, « Linked Data: The Story so Far », *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, 2011. [www.igi-global.com/chapter/linked-data-story-far/55046](http://www.igi-global.com/chapter/linked-data-story-far/55046) (consulté le mars 14, 2021).
- [9] K. Janowicz, P. Hitzler, B. Adams, D. Kolas, et C. Vardeman II, « Five stars of Linked Data vocabulary use », *Semantic Web*, vol. 5, n° 3, p. 173-176, 2014, doi: 10.3233/SW-140135.
- [10] A. Hasnain et D. Rebholz-Schuhmann, « Assessing FAIR Data Principles Against the 5-Star Open Data Principles », in *The Semantic Web: ESWC 2018 Satellite Events*, Cham, 2018, p. 469-477. doi: 10.1007/978-3-319-98192-5\_60.
- [11] D. Garijo et M. Poveda-Villalón, « Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web », *ArXiv*, 2020, doi: 10.3233/ssw200034.
- [12] N. Matentzoglou, J. Malone, C. Mungall, et R. Stevens, « MIRO: guidelines for minimum information for the reporting of an ontology », *J. Biomed. Semant.*, vol. 9, n° 1, p. 6, janv. 2018, doi: 10.1186/s13326-017-0172-7.
- [13] R. David *et al.*, « FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles », *Data Sci. J.*, vol. 19, n° 1, Art. n° 1, août 2020, doi: 10.5334/dsj-2020-032.
- [14] C. Bahim *et al.*, « The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments », *Data Sci. J.*, vol. 19, n° 1, Art. n° 1, oct. 2020, doi: 10.5334/dsj-2020-041.
- [15] M. D. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. Bonino da Silva Santos, et M. Dumontier, « A design framework and exemplar metrics for FAIRness », *Sci. Data*, vol. 5, juin 2018, doi: 10.1038/sdata.2018.118.
- [16] M. D. Wilkinson *et al.*, « Evaluating FAIR maturity through a scalable, automated, community-governed framework », *Sci. Data*, vol. 6, n° 1, Art. n° 1, sept. 2019, doi: 10.1038/s41597-019-0184-5.
- [17] « SurveyMonkey Powered Online Survey ». <https://www.surveymonkey.com/r/fairdat> (consulté le avr. 19, 2021).
- [18] M. Mokrane, L. Cepinskas, V. Åkerman, J. de Vries, et I. von Stein, « FAIR-Aware », 2020, Consulté le: mars 14, 2021. [En ligne]. Disponible sur: <https://pure.know.nl/portal/en/publications/fair-aware>
- [19] M. Poveda-Villalón, P. Espinoza-Arias, D. Garijo, et O. Corcho, « Coming to Terms with FAIR Ontologies », in *Knowledge Engineering and Knowledge Management*, Cham, 2020, p. 255-270. doi: 10.1007/978-3-030-61244-3\_18.
- [20] J. Frey, D. Streitmatter, F. Götz, S. Hellmann, et N. Arndt, « DBpedia Archivo: A Web-Scale Interface for Ontology Archiving Under Consumer-Oriented Aspects », in *Semantic Systems. In the Era of Knowledge Graphs*, vol. 12378, E. Blomqvist, P. Groth, V. de Boer, T. Pellegrini, M. Alam, T. Käfer, P. Kieseberg, S. Kirrane, A. Meroño-Peñuela, et H. J. Pandit, Éd. Cham: Springer International Publishing, 2020, p. 19-35. doi: 10.1007/978-3-030-59833-4\_2.
- [21] E. Amdouni et C. Jonquet, « FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies », avr. 2021. Consulté le: mai 20, 2021. [En ligne]. Disponible sur: <https://hal-lirmm.ccsd.cnrs.fr/lirmm-03208544>
- [22] N. F. Noy *et al.*, « BioPortal: ontologies and integrated data resources at the click of a mouse », *Nucleic Acids Res.*, vol. 37, n° Web Server issue, p. W170-W173, juill. 2009, doi: 10.1093/nar/gkp440.
- [23] S. Jupp, T. Burdett, C. Leroy, et H. Parkinson, « A new Ontology Lookup Service at EMBL-EBI », 2015.
- [24] C. Jonquet, A. Toulet, B. Dutta, et V. Emonet, « Harnessing the Power of Unified Metadata in an Ontology Repository: The Case of AgroPortal », *J. Data Semant.*, vol. 7, n° 4, p. 191-221, déc. 2018, doi: 10.1007/s13740-018-0091-5.
- [25] N. Juty, S. M. Wimalaratne, S. Soiland-Reyes, J. Kunze, C. A. Goble, et T. Clark, « Unique, Persistent, Resolvable: Identifiers as the Foundation of FAIR », *Data Intell.*, vol. 2, n° 1-2, p. 30-39, janv. 2020, doi: 10.1162/dint\_a\_00025.
- [26] C. Jonquet, A. Annane, K. Bouarech, V. Emonet, et S. Melzi, « SIFR BioPortal : Un portail ouvert et générique d'ontologies et de terminologies biomédicales françaises au service de l'annotation sémantique », Genève, Switzerland, juin 2016. Consulté le: mars 22, 2021. [En ligne]. Disponible sur: <https://hal.archives-ouvertes.fr/hal-01398250>