Project Title	**Expanding FAIR solutions across EOSC**
Project Acronym	**FAIR-IMPACT**
Grant Agreement No.	101057344
Start Date of Project	2022-06-01
Duration of Project	36 months
Project Website	https://fair-impact.eu

M4.4 - Review and analysis of Semantic Artefact Catalogues for serving FAIR semantic artefacts in EOSC

Work Package	WP4, Metadata and Ontologies
Lead Author (Org)	Clement Jonquet (INRAE), Nina Grau (INRAE)
Contributing Author(s) (Org)	Maria Poveda (UPM), Daniel Garijo (UPM), Vyacheslav Tykhonov (DANS-KNAW), Baptiste Cecconi (OBS-PARIS), Guillaume Alviset (CNRS/Data Terra), Ilaria Rosati (CNR), Martina Pulieri (CNR)
Due Date	2024-07-31
Date	2024-07-31
Version	V1.0
DOI	10.5281/zenodo.12799796
DOI of associated data	10.5281/zenodo.12799862

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Funded by the European Union

Versioning and contribution history

Version	Date	Author	Notes
0.1	03.07.2024	Nina Grau (INRAE), Clement Jonquet (INRAE)	First version of plan and structure of the document.
0.2	16.07.2024	Nina Grau (INRAE)	Preparation of content for Section 1 and first draft for Section 2.
0.3	18.07.2024	Nina Grau (INRAE)	Section 2 consolidated and first version of results in Section 3.
0.4	22.07.2024	Nina Grau (INRAE)	Commented results in Section 3.
0.5	23.07.2024	Clement Jonquet (INRAE)	Revised title, versioning. New Executive Summary. New Introduction.
0.6	23.07.2024	Clement Jonquet (INRAE)	Added new section 2 on related work. Section renumbering.
0.7	23.07.2024	Nina Grau (INRAE)	DOI reservation for report and associated data.
0.8	24.07.2024	Clement Jonquet (INRAE)	Revised Methodology (section 3.1).
0.9	25.07.2024	Clement Jonquet (INRAE)	Changed title. Finished Methodology with Section 3.1 and 3.2. Added Appendix 6.1.
0.10	29.07.2024	Clement Jonquet (INRAE) and Nina Grau (INRAE)	Added SAC technology table.
1.0	31.07.2024	Clement Jonquet (INRAE) and Nina Grau (INRAE)	Finished Section 4.2 and include Fig 7. Added conclusion. Final text reading and corrections.

Disclaimer

FAIR-IMPACT has received funding from the European Commission's Horizon Europe funding programme for research and innovation programme under the Grant Agreement no. 101057344. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

Table of Contents

Funded by the European Union

List of Tables

List of Figures

Funded by the European Union

Terminology

Acronym	Description
API	Application Programming Interface
CESSDA	Consortium of European Social Science Data Archives
EBI	European Bioinformatic Institute
EOSC	European Open Science Cloud
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DOI	Digital Object Identifier
FAIR	Findable, Accessible, Interoperable, Reusable principles
FAIR-enabling	Practices, methods or criteria that enhance the adoption or adherence of a digital object to the FAIR Principles
FFV	Features of a FAIR vocabulary
HTTP	HyperText Transfer Protocol
IF	Interoperability Framework
ISWC	International Semantic Web Conference
IVOA	International Virtual Observatory Alliance
LOV	Linked Open Vocabularies
MOD	Metadata for Ontology Description
NFDI	French National Research Institute for Agriculture, Food and Environment
OBO Foundry	Open Biological and Biomedical Ontology Foundry
OEG	Ontology Engineering Group
O'FAIRe	Ontology FAIRness Evaluator
OLS	Look Up Service
OORI	Open Ontology Repository Initiative
SA / SAs	Semantic Artefact
SAC /SACs	Semantic Artefact Catalogue
SKOS	Simple Knowledge Organisation System
SKOSMOS	Open source web-based SKOS vocabulary
SPARQL	SPARQL Protocol and RDF Query Language
SSHOC	Social Sciences & Humanities Open Cloud
SSSOM	Simple Standard for Sharing Ontological Mappings
TF	Task Force
URL	Uniform Resource Locator
WP	Work Package

Funded by the European Union

Executive Summary

In the rapidly evolving landscape of scientific research, the proliferation of ontologies and semantic artefacts necessitates the development of robust systems to manage and utilise these resources effectively. Semantic Artefact Catalogues (SAC) and ontology repositories are critical in this regard, especially within the framework of the European Open Science Cloud (EOSC) that has clearly identified the important role “ontologies and metadata” may have on the construction of a Web of FAIR data and services. These catalogues provide essential platforms for receiving, hosting, serving, aligning, and enabling the reuse of ontologies and other Semantic Artefacts (SA) (terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards). These catalogues not only facilitate the organisation and access of semantic artefacts but also support and sometime ensure their compliance with the FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles, which are foundational to the EOSC’s mission of promoting open science and data sharing across diverse scientific disciplines.

Semantic artefact catalogues essentially help their users to discover, manipulate, explore and exploit SAs without the need to manage or develop them. The types of SACs build by various (scientific but not only) communities ranges from simple semantic artefact listings to rich libraries with structured metadata, and advanced repositories (or portals) that offer a variety of services for multiple types of semantic artefacts,. These services may include browsing/searching, visualisation, metrics, recommendations, and annotation of data. SAC are often developed or maintained by specific discipline communities or infrastructures and we have seen the emergence of specific generic technologies –such as OntoPortal, SKOSMOS or OLS– that can be reused to deploy new semantic artefact catalogues.

Within FAIR-IMPACT’s WP4 on ontologies and metadata, T4.2 aims to establish guidelines and community practices with respect to the lifecycle of FAIR semantic artefacts from creation (T4.2.1) to sharing and reuse via catalogues or repositories (T4.2.2) and standardisation of SA metadata descriptions and SAC application programming interfaces. WP4 has already produced multiple deliverables showing the importance of SACs in the governance of semantic artefacts (M4.1, D4.1) and in their FAIR lifecycle (M4.2).

In this Milestone, we explore the current landscape of SACs, in EOSC and beyond; we make a quite comprehensive review of current and past SACs, sorting them by types, disciplines and technology. Plus, based on the five methodologies and tools for FAIRness assessment of SAs, available thru FAIR-IMPACT’s partners (O’FAIRe, FOOPS!, FsF, 10-SR and FVF) we have regrouped 10 important dimensions for FAIR semantic artefacts and we study how much each reviewed SAC enables or supports FAIR for their artefacts. The Milestone provides a good overview of how SACs can help SAs to address FAIR principles and contribute to the efficient management and utilisation of SAs.

The Milestone consists of the current report presenting our methodology and result analysis as well as associated data under the form of a spreadsheet which contains the listing of SACs, their classifications (status, type, discipline, technology) and the evaluation of their

Funded by the European Union

FAIR-enabling dimensions. The spreadsheet discussed and analysed in the current report is versioned with DOI: 10.5281/zenodo.12799862

Funded by the European Union

1. Introduction

1.1. FAIR Semantic Artefacts and FAIR-enabling SA-catalogues

The FAIR Data Principles[1] provides four core requirements to support the discovery, access and reuse of research data and have been largely promoted and adopted, especially in Europe and in the context of the deployment of the European Open Science Cloud (EOSC). We may cite initiatives and projects such as: GO FAIR[2], EC recommendations[3], FAIRsFAIR[4] or FAIRsharing[5]. Albeit these principles provide the "most basic levels of good data management and stewardship"[1] their implementation is not that simple. The notions of findability, accessibility, interoperability and reusability often depend on the digital object concerned, making different measuring variables and data sharing practices difficult to normalise. Indeed, while several recent reports and surveys[6,7,8] have revealed that even though the level of awareness of the FAIR Data Principles have increased, there is still a need for harmonisation of practices. FAIR-IMPACT has also argued[9] that FAIRness assessment shall be focused and specific to certain types of digital objects. **In this report, our digital objects of interest are Semantic Artefacts (SAs) i.e., a broader term to include ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards**. SA recognised as "*a machine-actionable and -readable formalisation of a conceptualisation enabling sharing and reuse by humans and machines*"[10], represent the highest level of meaningful knowledge representation within an interoperability framework, making them particularly sensitive to FAIR principles. Across all scientific disciplines, SAs are extensively utilised to represent and annotate data in a standardised way. These artefacts have become essential for adhering to the FAIR Data Principles, and they are increasingly recognized as research objects that must also comply with FAIR standards.

Data repositories play an important role in the development and dissemination of research outputs. Indeed, they often contribute to establishing standard data-sharing procedures within research communities. The inclusion (or deposit) of data in an *"open repository came out as the most important factor when determining the quality of a dataset*"[8]. And the role of data repositories to support FAIR (clearly established by Principle F4) is now considered

[1] Wilkinson M. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. https://doi.org/10.1038/sdata.2016.18
[2] GO FAIR. 2017. https://www.go-fair.org/go-fair-initiative/
[3] Turning FAIR into Reality (European Commission). 2018-2020. DOI https://doi.org/10.2777/1524
[4] FAIRsFAIR (European commission funded) . 2019. URL https://www.fairsfair.eu/the-project
[5] FAIRsharing.org. 2007. URL https://www.fairsfair.eu/the-project
[6] EOSC Task Force on FAIR metrics and Data quality's report. 2024. FAIR evaluation community survey. DOI: 10.5281/zenodo.10797764.
[7] Bahl R, et al. 2024. The Global Lens: Highlighting national nuances in researchers' attitudes to open data. DOI: https://doi.org/10.6084/m9.figshare.25569453.v1
[8] Science D, Hahnel M, Smith G. 2023. The State of Open Data 2023. DOI: https://doi.org/10.6084/m9.figshare.24428194.v1
[9] https://dx.doi.org/10.5281/zenodo.7848126
[10] Wim H., et al. 2020. D2.5 report of FAIR semantic recommendations second iteration. DOI: https://doi.org/10.5281/zenodo.4314320

Funded by the European Union

obvious. Beyond the fundamental use of an open and digital repositories to share datasets, multiples initiatives encourage this sharing process through quality and trustworthiness certifications, such as CoreTrustSteal[11] or ISO 16363 certifications[12]. **In this report, our data repositories of interest are Semantic Artefact Catalogues (SACs) i.e., a broader term to include libraries, registries, listing or repositories of semantic artefacts and also platforms often named terminology/vocabulary service/server.** Indeed, SAs are distributed in various formats, sizes, structures, and across overlapping domains, creating a need for unified platforms that can receive, host, serve, align, and enable their reuse in diverse communities and applications. The surge in the number of available ontologies and semantic artefacts has made SACs indispensable. SACs are typically designed to meet the specific needs of different communities. Their functionalities range from simple metadata listings, akin to libraries, to sophisticated platforms that provide advanced ontology-based services, such as browsing, searching, visualising, computing metrics, annotating and accessing data, recommendation of SA and assessing FAIRness, and sometimes even editing. Generally, SACs assist users in handling SAs without requiring them to manage the complex and time-consuming process of developing them. Additionally, like any other data repositories, they play a crucial role in making the SAs they host or serve FAIR. However, SACs are unequal with respect to how much they support or enable FAIR. Our goal is to establish a comparison framework to facilitate the selection and use of SACs in the context of EOSC and beyond.

1.2. FAIR-IMPACT project in the context of EOSC

1.2.1. EOSC environment

The implementation of the FAIR Data Principles is supported by the European Commission (EC) through the EOSC programme. The ambition of EOSC is to develop a "Web of FAIR Data and Services" for science in Europe, that breaks down the barriers of the compartmentalisation of research outcomes, tools, services and discipline spaces, enabling a multi-disciplinary environment. **Metadata and semantics are essential in the EOSC** for enhancing discoverability, interoperability, and efficient management of data. They enable seamless data integration and reuse across diverse research domains by using standardised vocabularies and ontologies, supporting automated processing, compliance, and collaboration. Rich metadata ensures data is FAIR aligning with open science principles. The **EOSC Semantic Interoperability Framework specifically highlights the importance of semantic artefact catalogues**, which are key components that facilitate the distribution, sharing and access of semantic artefacts.

1.2.2. FAIR-IMPACT work plan

The European Union's funded project FAIR-IMPACT supports the harmonisation and synchronisation of the FAIR enabling practices, aiming to realise a FAIR EOSC environment. **FAIR-IMPACT concentrates its efforts on FAIR-enabling practices and focuses on certain types of digital objects,** including datasets, research software, semantic artefacts and SA mappings. To address its objectives, one of the main goals of the FAIR-IMPACT project is to

[11] CoreTrustSteal repository certification. URL: https://www.coretrustseal.org/
[12] ISO 16363 repository certification. URL: http://www.iso16363.org/

Funded by the European Union

improve and promote the FAIRness of data and other digital objects by coordinating the implementation of frameworks of FAIR data practices at multiple levels and the EOSC environment. The FAIR-IMPACT project covers various components of the FAIR ecosystem, including persistent identifiers (WP3), **metadata and ontologies (WP4)**, metrics and certifications, and policies (WP5). The project provides guidelines for the EOSC to ensure these components are effectively integrated and utilised.

1.2.3. Scope of T4.2

WP4 entitled "Metadata and ontologies" aims to gather, synthesise and disseminate the materials needed to federate the approach to metadata and ontologies at various organisational and technical levels within the EOSC. Thus, WP4 is dealing with various types of research/digital objects including SAs and their mappings. WP4 has already produced several deliverables highlighting the importance of SACs in governing semantic artefacts (M4.1, D4.1) and ensuring their FAIR lifecycle (M4.2). Within WP4 on ontologies and metadata, **T4.2 focuses on establishing guidelines and community practices for the lifecycle of FAIR semantic artefacts**. This encompasses their FAIR-by-design creation (T4.2.1) and extends to their sharing and reuse through catalogues or repositories (T4.2.2), as well as the standardisation of SA metadata descriptions and application programming interfaces for SACs. Within T4.2, we work with multiple communities to consolidate, deploy or experiment semantic artefact catalogues for their discipline or context: AgroPortal (INRAE) for agri-food, and EcoPortal (LifeWatch) for ecology-biodiversity, EarthPortal (CNRS/ Data Terra) for earth sciences. OBS-PARIS is now also experimenting with the OntoPortal technology (https://ontoportal.org) (already used by Agro/Eco/EarthPortal) for the semantic artefacts in the astronomy area. With the participation of the OntoPortal Alliance, we are also reaching out to communities outside of FAIR-IMPACT, (e.g., NFDI4Biodiversity in Germany). This joint work on semantic artefact catalogues has enabled, in partnership with T5.3, to make the O'FAIRe ontology FAIRness assessment tool available in EcoPortal and EarthPortal in addition to AgroPortal; this tool was then used in the FAIRness challenge support action. This is a good example of FAIR-IMPACT's mission to take FAIR enabling tools and methods from one community to the other. T4.2 is also consolidating the MOD (*Metadata for Ontology Description and Publication Ontology*) to specify a DCAT2-based standard way to describe semantic artefacts. We have released several versions (https://github.com/FAIR-IMPACT/MOD) of our specification (MOD v3.2 as latest) for semantic artefact description (M4.3), captured mappings between metadata vocabularies used within MOD using the SSSOM format (https://github.com/FAIR-IMPACT/MOD-mappings) and we are preparing a standard API for semantic artefact catalogues based on MOD (https://github.com/FAIR-IMPACT/MOD-API) that will be the subject of the D4.3.

1.2.4. Position with respect to the Semantic Interoperability Task Force work

In parallel, the Semantic Interoperability Task Force (TF) (finished end of 2023) also had a dedicated "topic" on semantic artefact and did produce results related to SACs:[13]

[13] https://doi.org/10.5281/zenodo.10843882

Funded by the European Union

- The TF recognised "the Semantic Artefact Catalogue component described in the EOSC Interoperability Framework (IF) as a critical part of the long-term viability of any research data infrastructure" in one of their 5 broad recommendations;
- The TF emphasised the importance of assessing the maturity of semantic artefact catalogues with **a maturity model developed to evaluate these catalogues**, providing governance recommendations and addressing interoperability challenges. This model includes twelve dimensions, tested on 26 catalogues, to specify levels of compliance and maturity. It targets semantic artefact providers, users, and catalogue developers, offering criteria to enhance their resources;
- The TF published the maturity model of SACs as a journal paper.[14]

Our work is complementary to the task force effort in two main aspects: (i) our review of SACs is larger than the TF's and targets to be historical and comprehensive and include various types of catalogues; (ii) our assessment is primarily focused on FAIR-enabling capabilities. Still, some members of FAIR-IMPACT's WP4 were associated with the TF work as TF members. C. Jonquet (lead of T4.2 but not a TF member), did participate in the TF report on maturity of SACs to ensure the convergence and cross-fertilization of the work which occurred within FAIR-IMPACT and within the TF in parallel.

1.2.5. Milestone description

T4.2's objectives include supporting the establishment and/or development of SACs and fostering their coordination within the EOSC, thereby promoting FAIR enabling practices for semantic artefacts. Among the initial steps of this promotion of SAC, FAIR-IMPACT included reviewing SACs and comparing the technologies currently used by a broad spectrum of SAC within EOSC and beyond. **This milestone delineates the collaborative work undertaken by T4.2 to achieve the task's objectives of gathering, reviewing, and analysing as much SAC currently available or that previously exist.** We examine the current landscape of SACs within EOSC and beyond, providing a comprehensive review of both current and past SACs categorised by types, disciplines, and technologies. Additionally, using five methodologies and tools for FAIRness assessment of semantic artefacts available through FAIR-IMPACT's partners (O'FAIRe, FOOPS!, FsF, 10-SR, and FVF), we have identified 10 key dimensions for FAIR semantic artefacts. We analyse how each reviewed SAC supports or enables these FAIR dimensions. This milestone offers an insightful overview of how SACs facilitate the adherence of semantic artefacts to FAIR principles, contributing to their effective management and utilisation.

The report is structured into four comprehensive sections: Section 1 introduces the overarching context and complexities of SA and SAC, and how the EOSC environment and FAIR-IMPACT project integrate them into their objectives. Section 2 redraws the academic work related to SAC and the technologies they employ. Section 3 outlines the methodology used for a comprehensive review of SAC and describes how we classified SAC based on their status, types, technology used and FAIR-enabling dimensions used to evaluate how integrating SA into SAC can enhance their FAIRness. Each FAIR-enabling dimension for SACs is presented. Section 4 presents our results and analysis, with a state-of-the-art list of SACs

[14] https://dx.doi.org/10.1038/s41597-024-03185-4

Funded by the European Union

gathered, reviewed and evaluated against the FAIR-enabling dimensions also presented in this work.
The milestone consists of the current report presenting our methodology and result analysis as well as **associated data under the form of a spreadsheet which contains the listing of SACs, their classifications (status, type, discipline, technology) and the evaluation of their FAIR-enabling dimensions**. The spreadsheet discussed and analysed in the current report is versioned with DOI: 10.5281/zenodo.12799862

2. Related work on semantic artefact catalogues

In 2023, we published an article at the *22nd International Semantic Web Conference (ISWC)* related to the OntoPortal technology largely used within FAIR-IMPACT.[15] This article included a section on historical perspective on ontology repositories and SACs that we reproduce as this here. References of this section can be found in Section 7.

2.1. From ontology libraries and repositories to semantic artefact catalogues

With the growing number of developed ontologies, ontology libraries and repositories have been a long-time interest in the semantic web community. Ding & Fensel [11] presented in 2001 a review of **ontology libraries:** "*A system that offers various functions for managing, adapting and standardising groups of ontologies. It should fulfil the needs for re-use of ontologies.*" Ontology libraries usually register ontologies and provide metadata description. The terms collection, listing or registry were also later used to describe similar concepts to ontology libraries. All correspond to systems that help reuse or find ontologies by simply listing them (e.g., DAML, Protégé or DERI listings) or by offering structured metadata to describe them (e.g., FAIRSharing, BARTOC, Agrisemantics Map). But those systems do not support any services beyond description, including services based on the content of the ontologies. In the biomedical domain, the OBO Foundry [12] is a reference library effort to help the biomedical and biological communities build their ontologies with an enforcement of design and reuse principles. A number of services and tools are built to work with this library of semantic artefacts.
Hartman et al. [13] introduced in 2009 the concept of **ontology repository**: "*A structured collection of ontologies (…) by using an Ontology Metadata Vocabulary. References and relations between ontologies and their modules build the semantic model of an ontology repository. Access to resources is realised through semantically-enabled interfaces applicable for humans and machines.*". Multiple ontologies repositories have been developed since then, with advanced features such as search, metadata management, visualisation, personalization, mappings, annotation and recommendation services, as well as application programming interfaces to query their content/services. Here again the biomedical domain has seen a lot of resources (not necessarily synchronised), such as the NCBO BioPortal [8], OntoBee [14], the EBI Ontology Lookup Service (OLS) [15] and AberOWL [16]. We have seen also repository initiatives such as the Linked Open Vocabularies [17], OntoHub [18], and the

[15] https://dx.doi.org/10.1007/978-3-031-47243-5_3

Funded by the European Union

Marine Metadata Initiative's Ontology Registry and Repository [19] and its earth science counterpart, the ESIP Federation's Community Ontology Repository. By the end of the 2000's, the topic was of high interest as illustrated by the 2010 ORES workshop [20] and the 2008 Ontology Summit.[16] More recently, the SIFR BioPortal [21] prototype was built to develop a French Annotator and experiment with multilingual issues in BioPortal [22]. The first reuse of the OntoPortal technology to develop a free and open, community-driven ontology repository in the spirit of BioPortal, but for agri-food, was AgroPortal, started at the end of 2015 [23]. D'Aquin & Noy [24] and Naskar and Dutta [25] provided the latest reviews of ontology repositories.

In parallel, there have been efforts to index any semantic web data online (including ontologies) and offer search engines such as Swoogle and Watson [26, 27]. We cannot consider these **"semantic web indexes"** as ontology libraries, even if they support some features of ontology repositories (e.g., search). Other similar products are terminology services or vocabulary servers which are usually developed to host one or a few terminologies for a specific community (e.g., SNOMED-CT terminology server, UMLS-KS, CLARIN vocabulary services, OpenTheso, etc); they are usually not semantic web compliant and did not handle the complexity of ontologies, although an increasing number of terminology services are getting compliant with SKOS (Simple Knowledge Organization System) [28]. We can also cite the ARDC Research Vocabularies Australia (https://vocabs.ardc.edu.au) using multiple technologies such as PoolParty and SSISVoc.

In the following, we will focus on ontology repositories considering they offer both ontology-focused services (i.e., services for ontologies) and ontology-based services (i.e., services using ontologies). We will also name them now **semantic artefact catalogues**, a term which emerged in the forum and discussions around building the EOSC (e.g., [29]) and which translates the idea that such catalogues are not only for ontologies but must offer common services for a wide range of semantic artefacts.

2.2. *Generic ontology repository and semantic artefact catalogue technology*

In the end of the 2000's, the Open Ontology Repository Initiative (OORI) [30] was a collaborative effort to develop a federated infrastructure of ontology repositories. At that time, the effort already reused the NCBO BioPortal technology [31] that was the most advanced open-source technology for managing ontologies at that time. Later, the initiative studied OntoHub [18] technology for generalisation but the Initiative is now discontinued.

In the context of our projects, to avoid building new ontology repositories from scratch, most of the authors have considered which of the technologies cited above were reusable. While there is a strong difference between "open source" (most of them are) and "made to be reused" we think only the NCBO BioPortal and OLS were really generic ontology repository candidates for both their construction and documentation. OLS technology has always been open source but some significant changes (e.g., the parsing of OWL) facilitating the reuse of the technology for other portals were done with OLS 3.0 released in December 2015. Until very recently (2022), in the context of the NFDI projects

[16] http://ontolog.cim3.net/wiki/OntologySummit2008.html

Funded by the European Union

(https://terminology.tib.eu), we had not seen another public repository built with OLS. On the other hand, the NCBO BioPortal was developed from scratch as a domain-independent and open-source software. Although it has been very early reused by ad-hoc projects (e.g., at OORI, NCI, and MMI), it is only in 2012, with the release of BioPortal 4.0 that the technology, made of multiple various components was packaged as a virtual appliance, a virtual server machine embedding the complete code and deployment environment, allowing anyone to set up a local ontology repository and customise it. The technology is denoted as OntoPortal since 2018.
Skosmos [32] is another alternative originally built in for reuse, but it only supports browsing and search for SKOS vocabularies. For instance Finto (https://finto.fi) or Loterre (www.loterre.fr) have adopted Skosmos as backend technology. Another example is VocPrez, an open-source technology developed by a company adopted for examples by the Geoscience Australia Vocabularies system (https://vocabs.ga.gov.au) or by the NERC Vocabulary Server (http://vocab.nerc.ac.uk). Another technology is ShowVoc, based on the same technological core as VocBench but it appears to have drawn inspiration from OntoPortal in terms of its design and services.

3. Methodology

3.1. Comprehensive listing of Semantic Artefact Catalogues

3.1.1. Related work and listing process

Each T4.2 member (a dozen of people) was initially asked to list any SAC they were aware of, and/or any resources listing a collection of SAC, to a shared document. The list was first populated with resources identified within:

- C. Jonquet's Habilitation Background chapter;[17]
- Noy and d'Aquin 2012's historical article;[18]
- Naskar and Dutta's article based on an internal ISI report in 2016;[19]
- The recent work of the EOSC Semantic Interoperability TF theme 2 (to which C. Jonquet was associated) on establishing a maturity model for SAC;[20]
- The ISWC 2023 article cited Section 2 presenting the OntoPortal technology;[21]
- The BARTOC Terminology Registries list which did a similar exercise. Approximately 30% of the individual SAC gathered during our review were sourced from this list (i.e., not found on any other previously listed sources), which significantly enriched our SAC collection.

[17] Clement Jonquet. Ontology Repository and Ontology-Based Services – Challenges, contributions and applications to biomedicine & agronomy. Web. Université de Montpellier, 2019. ⟨tel-02133335⟩
[18] https://doi.org/10.1016/j.websem.2011.08.005
[19] https://www.isibang.ac.in/~bisu/paper/ETD_OntologyLibraries-A%20Study%20from%20an%20Ontofier%20and%20an%20Ontologist%20Perspectives.pdf
[20] https://doi.org/10.1038/s41597-024-03185-4
[21] https://dx.doi.org/10.1007/978-3-031-47243-5_3

Funded by the European Union

Then we also reviewed the homepages of some well known SAC technology (OntoPortal, SKOSMOS, TemaTres, OpenTheso) which listed the installations of their technology. In addition, we have explicitly contacted the providers of each of the technologies, already known or identified during our study, asking them for an exhaustive listing of the installations of their technologies, if such a listing was available and public.[22]

Once all individual SAC were roughly identified, two members of T4.2 (N. Grau, C. Jonquet) transferred them into a shared working spreadsheet comprising 10 properties listed in Table 1 below. Once all fields of the table were completed, we performed a double-checking to verify the information, complete missing information and harmonised the whole content. Members of T4.2 contributed to the table too, especially for the SAC they were very familiar with or responsible for.

Table 1. List of the properties gathered for each SAC listed into the shared spreadsheet.

Name	URL	Discipline	Type	Status	Generic technology	Contact	Support organisation/project	Description	Description source

The properties about each SAC were filled-in as explained in the following sections. They will be used in Section 4 for the analysis.

3.1.2. General information

For each SAC reviewed, we identified the name and current Web page URL (Uniform Resource Locator) hosting the catalogue online. We also found a contact and supporting organisation or project as well as a short “description” usually taken from the catalogue itself or another related resource (recorded too).

3.1.3. Status

The goal of the FAIR-IMPACT review of semantic artefact catalogues was of course to establish a state-of-the-art listing of what catalogues are available now for use. However, we also believe that when we knew about an existing relevant catalogue in the past, now discontinued, it would make sense to include it in our review. In this sense, our review also plays the role of an archive.
The status of the SAC is mostly determined based on the maintenance/availability of its URL. Thus, a SAC is considered "Retired" if the URL is no longer accessible and/or functional, or if there has been no activity on the website for more than several years. Conversely, the SAC's status is considered "Active". Otherwise the “Prototype” status is referring to catalogues that we know are under construction and did not achieve their final form or role yet.

3.1.4. Types of Semantic Artefact Catalogues

In this report, we define a SAC as a Web platform where multiple (more than one) SA are at minimum listed or indexed, but in most cases stored, preserved, served and shared. The expression SAC is therefore a broader term to include multiple historical/legacy terms used in the literature such as libraries, registries, listing, repositories or even service or server.

[22] To the exception of Centre (a private solution) and OCLC that we considered obsolete and were reported too late in our study.

Funded by the European Union

While SAC platforms are conceptually similar, they can differ according to their architecture and/or the type of services they offer. During our review, we have regrouped catalogues in the following types:

- Listing: This is a simple online listing of semantic artefacts on a web page or web site. It consists of a continuous series of SAs, ordered or not. Usually, only the basic information about an SA is provided such as a name and a link to where to find/access the SA.
- Library: This is a structured online listing of semantic artefacts with rich harmonised metadata but without hosting or serving the content of artefacts. Libraries may offer various functions for identifying, managing, grouping semantic artefacts. They are usually provided by a small group or specific community which decides about the inclusion of the SA within the library based on some criteria. Library applications are almost always based on some ad-hoc technology.
- Search service: This is an online index that is specifically dedicated to indexing/searching semantic artefacts. It does not host or serve directly the content of artefacts, but in a search engine approach, it offers means to identify SA content and link to it.
- Repository: This is an advanced web application offering the features of a library (harmonised rich metadata) but also hosting and serving the content of semantic artefacts. It provides advanced features to search (similar to Search service), browse, manage metadata, and sometimes additional related services such as mapping hosting and automatic generation, text annotation or recommendation of SAs. Repositories can be queried by machine via APIs or directly SPARQL endpoints; they are often based on a common, generic technology.
- Other: This is a SAC that does not fit with the previously defined types but that deserves to be listed. In our report, we identified only three SACs as “Other”.

3.1.5. Disciplines

For each SAC, we tried to determine if it is focused on a specific domain or discipline. The assignment of a discipline was done in two steps.

The first step was conducted by two members of T4.2 who categorised each SAC with “open” keywords corresponding to disciplines, based on their knowledge of the listed SACs, their content, and the descriptions provided on the respective websites. This process yielded 38 unharmonized disciplines, with a 10% representation of unique disciplines. And 85 SACs were classified as “General” when they were not specific to a domain or we could not easily tell which one.

In a second step, to harmonise the (non General) disciplines, in order to ease comparison and analysis, we decided to align them with already established domain’s classifications. In our case, we used the list of academic disciplines provided by Wikipedia.[23] To achieve this alignment, we used a large language model (i.e., ChatGPT)[24] to classify the SAC from our collection based on their Web URL. The prompt used was: "*According to the list proposed by Wikipedia: https://en.wikipedia.org/wiki/List_of_academic_fields, classify by field of*

[23] Wikipedia of academic fields. URL: https://en.wikipedia.org/wiki/List_of_academic_fields#
[24] OpenAI. "ChatGPT-4." 2024. URL: https://chatgpt.com/

Funded by the European Union

academic discipline all the following sites according to their content while specifying to which sub-domain they belong." Subsequently, we compared and corroborated manually the two sets of classifications (provided by the T4.2 members and ChatGPT). We refined the classification to 17 disciplines standardised in Wikipedia, with less than 5% representation of unique disciplines. The description of each of the disciplines is available on the corresponding Wikipedia webpage.[16] We shall use only these 17 disciplines plus the "General" category in our analysis next Section.

3.1.6. Generic technology

As much as possible, we tried to identify if the SAC was running some kind of generic technology that is (or could be) used to deploy several catalogues (e.g., in different projects or disciplines). The SAC technology was determined by our knowledge of this area and acquaintance with common technologies (there is not so much) available to deploy or install a new SAC. Typically, a technology used to set up a SAC was considered "generic" if we could find a description of this technology on its own (typically a web site and/or a source code repository) or if we could find another SAC running the same technology.
The different SAC generic technology identified have been listed apart[25] and for each of them we identified (Table 2): an URL of a landing page presenting the technology, if the code was open source or not and if yes, the availability of the source code repository, the type of SAC the technology allow to implement (in all cases this was Repository), the current status of the technology (active or retired), the contact details, the supporting organisation, and a short description (with source).

[25] The list of SAC technology is also available in a dedicated tab in the spreadsheet associated with this report.

Funded by the European Union

Table 2. List of Semantic Artefact Catalogue technologies.

Name	Related URL	Open source (Yes/No)	Code repository	Type of SAC supported	Status	Contact
OntoPortal	https://ontoportal.org/	Yes	https://github.com/ontoportal	Repository	Active	Clement Jonquet
	The OntoPortal Alliance (https://ontoportal.org) is a consortium of several research and infrastructure teams and a company dedicated to promoting the development of ontology repositories—in science and other disciplines—based on the open, collaboratively developed OntoPortal open-source software. Teams in the Alliance develop and maintain several openly accessible ontology repositories and semantic artefact catalogues. These ontology repositories include BioPortal, the primary and historical source of OntoPortal code, but also AgroPortal, EcoPortal, MatPortal and more. The OntoPortal Alliance's original motivation and vision was to reuse outcomes and experiences obtained in the biomedical domain—an area where the use of ontologies has always been important—to serve and advance other scientific disciplines. (Source: https://dx.doi.org/10.1007/978-3-031-47243-5_3)					
OLS based	https://github.com/EBISPOT/ols4	Yes	https://github.com/EBISPOT/ols4	Repository	Active	Henriette Harmse
	The Ontology Lookup Service (OLS) is a repository for biomedical ontologies that aims to provide a single point of access to the latest ontology versions. You can browse the ontologies through the website as well as programmatically via the OLS API. OLS is developed and maintained by the Samples, Phenotypes and Ontologies Team (SPOT) at EMBL-EBI. (Source: https://www.ebi.ac.uk/ols4)					
SKOSMOS	https://skosmos.org/	Yes	https://github.com/NatLibFi/Skosmos	Repository	Active	Osma Suominen
	Open source web-based SKOS browser and publishing tool. Features; search and browse vocabularies, alphabetical index, thematic index, structured concept display, visualized concept hierarchy, multilingual user interface. Access to controlled vocabularies for indexing, information retrieval and vocabulary development. (Source: https://skosmos.org/)					
ShowVoc	https://showvoc.uniroma2.it/	Yes	https://bitbucket.org/art-uniroma2/showvoc/src/master/	Repository	Active	Armando Stellato
	ShowVoc is a web-based, multilingual, platform for publishing and consulting OWL ontologies, SKOS(/XL) thesauri, Ontolex-lemon lexicons and generic RDF datasets.ShowVoc business and data access layers are realized on top of Semantic Turkey, an open-source platform for Knowledge Acquisition and Management realized by the ART Research Group at the University of Rome Tor Vergata. ShowVoc offers a powerful browsing environment, with facilities for inspecting OWL ontologies, SKOS/SKOS-XL thesauri, OntoLex lexicons and any sort of RDF dataset. Cross-dataset features, such as global search and the translation API benefit from the presence of different datasets in order to realize a multilingual resource for term reference and authoritative term translation. (Source: https://showvoc.uniroma2.it/)					
TemaTres	https://vocabularyserver.com/web/about	Yes	https://github.com/tematres/TemaTres-Vocabulary-Server/	Repository	Active	Diego Ferreyra
	TemaTres is an open source vocabulary server, web application to manage and exploit vocabularies, thesauri, taxonomies and formal representations of knowledge. (Source: https://vocabularyserver.com/web/about)					
ONKI SKOS Server	https://seco.cs.aalto.fi/services/onkiskos/	Yes	https://seco.cs.aalto.fi/services/onkiskos/onki-skos-20121221.zip	Repository	Retired	Jouni Tuominen
	ONKI SKOS is a server for lightweight vocabularies in SKOS and ontologies in RDFS/OWL format. Using ONKI SKOS, a vocabulary with related AJAX mash-up and Web Service support can be published and used in applications cost-efficiently with very little extra work. (Source: https://seco.cs.aalto.fi/services/onkiskos/)					
Vocab	https://github.com/oeg-upm/vocab.linkeddata.es	Yes	https://github.com/oeg-upm/vocab.linkeddata.es	Repository	Active	María Poveda / Daniel Garijo
	This repository contains the source code for generating the website published at http://vocab.linkeddata.es. If you want to add a new vocabulary to the site you only need to include its URI in the Vocabularies CSV. (Source: https://github.com/oeg-upm/vocab.linkeddata.es)					
OpenTheso	https://opentheso.hypotheses.org/	Yes	https://github.com/miledrousset/Opentheso2	Repository	Active	Miled Rousset

Funded by the European Union

	Opentheso is a multilingual and multi-hierarchical thesaurus manager. It complies with ISO 25964-1:2011 and ISO 25964-2:2012 (Information and documentation. Thesaurus and interoperability with other vocabularies). It is distributed as open source under the CeCILL_C license, a free French law license compatible with the GNU GPL license.Created in 2005 at the request of the Federation and Resources on Antiquity (GDS Frantiq), for the management of the Pactols thesaurus, it is today positioned as a generic tool offered in the TGIR Huma-Num service grid. It is developed under the direction of Miled Rousset, head of the Web Semantics and Thesauri (WST) technological platform at the Maison de l'Orient et de la Méditerranée (MOM) and IT director of GDS Frantiq. (Source :https://opentheso.hypotheses.org/)					
iQvoc	https://iqvoc.net/	Yes	https://github.com/innoq/iqvoc	Repository	Active	info@innoq.com
	iQvoc supports vocabularies that are common to many knowledge organisation systems, such as: Thesauri, Taxonomies, Classification schemes, Subject heading systems. (Source: https://iqvoc.net/)					
ORR	https://mmisw.org/orrdoc/about/	Yes	https://github.com/mmisw	Repository	Active	Carlos Rueda
	Developed by MMI1, the Ontology Registry and Repository (ORR) is a web application and service to create, update, access, and map ontologies and their terms. Funding for the version 3 modifications to the ORR software were provided by the EarthCube X-DOMES Project. (Source: https://mmisw.org/orrdoc/about/)					
VocPrez	https://rdflib.dev/VocPrez/	Yes	https://rdflib.dev/VocPrez/	Repository	Retired	Nicholas Car
	A read-only web delivery system for Simple Knowledge Organization System (SKOS)-formulated RDF vocabularies. (Source: https://rdflib.dev/VocPrez/)					
Centree	https://scibite.com/platform/centree-ontology-management-platform/	No	NA	Repository	Active	Simon Jupp
	SciBite's ontology management platform CENtree provides a centralised, enterprise-ready resource for ontology management and transforms the experience of maintaining and releasing ontologies for research-led businesses. CENtree combines ease of use with cutting-edge artificial intelligence techniques to assist users, for example, by suggesting possible relationship connections for a given ontology class. (Source: https://scibite.com/platform/centree-ontology-management-platform/)					
Termweb	https://www.interverbumtech.com/products-services/termweb/	No	NA	Repository	Active	info.eu@interverbumtech.com
	There are many things that make a great terminology management platform. It must be powerful—with updated processing capabilities that allow it to quickly handle complex data at high volumes. And it should be inherently versatile: full of new features that lets users work the way they want to work, from any location and in any environment. But most of all, it should be user-friendly, so that anyone in your organization can feel comfortable using it. That's what makes TermWeb 4 the most advanced terminology management software available today. (Source: https://www.interverbumtech.com/products-services/termweb/)					
Data Harmony Taxonomy Suite	https://www.accessinn.com/data-harmony-products/	No	NA	Repository	Active	Contact form
	We are experts in the design and delivery of semantic solutions. Founded in 1978 and headquartered in Albuquerque, NM, Access Innovations is a pioneer and the industry leader in making information assets easily and automatically discoverable. Our major advantage is our patented, award-winning Data Harmony software combined with the technical expertise to create information discovery systems that meet unique client needs. Access Innovations has built thousands of controlled vocabularies through more than 2,000 client engagements. We deliver clean, well-formed, metadata-enriched content so our clients can reuse, repurpose, store, and find their knowledge assets. We go beyond the standards to build taxonomies and other data control structures as a solid foundation for your information. (Source: https://www.accessinn.com/about-us/)					
OCLC Terminology Services	http://tspilot.oclc.org/resources/	No	NA	Repository	Retired	oclc@oclc.org
	The Terminology Services prototype uses library and web standards to make the terms, relationships, descriptions, and other information in controlled vocabularies available as resources on the web. (Source: https://www.oclc.org/content/dam/research/activities/termservices/resources/termservices-overview.pdf)					

Funded by the European Union

3.1.7. Exclusion criteria

We decide to not include in the list of SACs the tool, website or application (vocabulary services or terminology service) with only one semantic artefact. Indeed, many ad-hoc web applications were developed to host and served on unique semantic artefact and we estimate the rationale behind the existence of such servers or service was very much different from the ones of building a catalogue of multiple semantic artefacts to list or serve them to a specific community. While these "SAC with only one SA" have to be excluded from our study, we decided to keep some of them encountered during our systematic review. Similarly, we also encountered other types of resources or platforms related to vocabularies, ontology engineering, metadata or semantic artefacts in general and kept a few of these resources too.[26]

3.2. FAIR-enabling dimensions

3.2.1. From SA FAIRness assessment to SAC FAIR-enabling

Our second objective was to analyse or assess how much a SAC would support or help the artefact hosted/served adhere to the FAIR Principles. In FAIR-IMPACT, we call this "FAIR-enabling" i.e., the capacity of a tool, method or software to help make things (data or other digital objects) FAIR.

Within FAIR-IMPACT's WP4, we gather members that have produced all the state-of-the-art methods or tools to evaluate/assess the level of FAIRness of semantic artefacts. These tools and methods are:

- **O'FAIRe (Ontology FAIRness Evaluator)**[27] a methodology and tool designed to assess the FAIRness of ontologies and semantic artefacts, based 61 questions, with 80% of the assessment relying on resource metadata descriptions. Originally implemented within the AgroPortal semantic artefact catalogue, and later transferred to multiple other SACs in the context of FAIR-IMPACT, O'FAIRe provides both global and detailed scores, helping users visualise and improve the FAIRness of their resources through user-friendly interfaces like the FAIRness wheel.
- **FOOPS! (Ontology Pitfall Scanner for the FAIR principles)**[28] "a web service designed to assess the compliance of vocabularies or ontologies against the FAIR principles. FOOPS! performs a total of 24 different checks from the four FAIR dimensions, reflecting the best practices and latest community discussions to adapt FAIR to semantic artefacts. The web service not only detect best practices according to each principle, but also offers an explanation of why a particular principle fails, and helpful suggestions to overcome common issues."
- **FAIRsFAIR's Recommendations (FsF Precs)**[29] were produced in March 2020, within the eponym H2020 project. It's a list of 17 recommendations and 10 best practices

[26] Some examples of "SAC with only one SA" and resources "Other not SAC" are available in a dedicated tab in the spreadsheet associated with this report.
[27] https://dx.doi.org/10.1504/IJMSO.2022.131133
[28] http://ceur-ws.org/Vol-2980/paper321.pdf
[29] https://doi.org/10.5281/zenodo.5362010

Funded by the European Union

recommendations for making semantic artefacts FAIR. 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.

- **10-simple rules (Ten simple rules for making a vocabulary FAIR)**[30] presented in a paper which outlines ten rules for converting a vocabulary into a FAIR vocabulary. The guidelines cover vocabulary and term metadata, development, and maintenance. Following these rules ensures the vocabulary can be used for unambiguous data annotation, enhancing data interoperability and integration.
- **FFV (Features of a FAIR vocabulary)**[31] presented in a paper which proposes FAIR Vocabulary Features with existing indicators, and demonstrated using biomedical vocabularies. The conclusions provide features and indicators for assessing FAIR vocabularies, identify use cases for vocabulary engineers, and offer guidance for vocabulary development and improvement.

These methods & tools are specific to semantic artefacts; they often identify the importance of hosting the artefacts in a relevant catalogue as an important aspect of enabling FAIR. In some cases, some criterias of these tools and methods are directly depending on the catalogues and not on the artefact itself. For example, O'FAIRe contains multiple FAIRness assessment questions –e.g., related to technical accessibility– that are "automatically addressed" by the catalogue on which O'FAIRe is evaluating FAIRness of an artefact. As another example, some recommendations of FAIRsFAIR are directly addressed to the catalogues (PRec 5 or 7).

We reviewed the criterias, questions or recommendations of the 5 tools and methods listed above and regrouped them in 10 broad dimensions. This is not an alignment or uniformisation of the 5 tools and methods which each keep their diversity and specificity in assessing FAIRness but this more of a broader prism to uniformly see them. The 10 broad dimensions are:

- **Identifiers.** Include criterias related to SA's identifiers to ensure global uniqueness, persistence, and resolvability. The presence of external identifiers such as DOIs. Metadata identifiers, version-specific URIs, and unique identifiers for SA versions to guarantee traceability, versioning and accessibility. Each term/object within a SA also receives a unique identifier, ensuring comprehensive and reliable identification.
- **Metadata**. Include criterias related to SA metadata e.g., follows the MIRO guidelines, including mandatory properties like title, description, licence, and creation date. Additional metadata properties are encouraged too. Metadata can be included within the SA source file or described externally with clear linking. Rich metadata for SA and SA-content enhances usability, with information about class/concepts,

[30] https://doi.org/10.1371/journal.pcbi.1009041
[31] https://doi.org/10.1186/s13326-023-00286-8

Funded by the European Union

properties, definitions, hierarchy, labels, etc. Provenance information and a consistent schema are essential for maintaining the SA's integrity and usability.

- **Cataloguing**. Include criterias related to public registration of the SA in multiple libraries and repositories, ensuring wide accessibility and indexing by web search engines. Public registration in trustworthy repositories enhances discoverability and maintenance traceability. Indexed SA are easier to search, find, and use, promoting broader adoption and interoperability.
- **Resolvability**. Include criterias related to the accessing and resolvability of URIs and identifiers (which should resolve to the SA or SA-content itself and support content negotiation). Resolvability ensures that users can access the SA and its metadata reliably. Community standard APIs and multiple serialisations formats for machine readability. Ensuring that both humans and machines can access SA and SA-content enhances usability and integration into various systems.
- **Access protocols**. Include criterias related to the use of standardised, open, free, and universally implementable protocols such as HTTP/HTTPS for SA access. SPARQL endpoints and other protocols should also be supported if they meet the criteria of openness and universal implementation. Secure access protocols with authentication and authorization capabilities ensure that sensitive information is protected. Accessibility through common protocols facilitates the integration of the ontology into diverse applications, maintaining open and controlled access as needed.
- **Versioning**. Include criterias related to hosting SA in repositories that support versioning, with metadata available for each version. Metadata should remain accessible even if the SA itself is no longer available. Clearly informing the status and changes of each version supports transparency and traceability. Appropriately versioned SA ensures that users can track the evolution and usage of the SA over time, providing a stable foundation for continued development and use.
- **Encoding**. Include criterias related to the encoding of the SA that must use standardised representation languages, preferably those recommended by the W3C. The representation language, syntax, and formality level should be clearly informed. Availability in multiple syntaxes/formats ensures broader applicability and machine readability. Using formal, accessible languages promotes interoperability and integration into various semantic web technologies, supporting the SA functionality.
- **Reuse, mappings & crossrefs**. Include criterias related to SA importing and reusing terms from other FAIR SAs, ensuring qualified references and alignments are well represented and curated. Information about the influence of other SAs and clear mappings between SA should be provided. Standard vocabularies should be used for describing SA metadata. Documenting crosswalks, mappings, and reuses between SAs facilitates integration, enhancing the SA's interoperability and reuse potential.
- **Licensing**. Include criterias related to SA's licence access rights, permissions, usage guidelines, and copyright holders. Ensuring that the licence is machine-readable.. Open and clear licensing arrangements foster broader adoption and compliance with legal standards, supporting the SA's sustainable and ethical use.
- **Provenance**. Include criterias related to SA provenance information, including actors involved, accrual methods, versioning, latest changes, methodology, tools, and

Funded by the European Union

funding organisations. Clear documentation, supports transparency and reproducibility. Provenance ensures that the SA's development and updates are traceable, which is crucial for trust and reliability. Meeting community standards and including endorsements by projects or organisations enhances credibility and adoption within specific domains.

Appendix 6.1 lists all the criterias, questions or recommendations in each dimension.

3.2.2. FAIR-enabling levels

For each SAC identified, we asked ourselves how much the SAC supports or enables the criterias within a dimension. In other words, to which level including an semantic artefact within a catalogue helps this SA becomes more FAIR. We fixed 3 FAIR-enabling levels as described in Table 3.

Table 3. FAIR-enabling level for Semantic Artefact Catalogues with respect to the FAIR-enabling dimensions.

	The SAC mostly handles this dimension (either automatically or through manual check). SA developers do not have anything else to do.
	The SAC helps address the dimension (seen as a requirement to get in the SAC) but does not fully handle it. SA developers may have to take care of addressing technical aspects of these criterias.
	The SAC mostly does not support this dimension at all. It's totally independent of the SAC for SA developers to address these criterias.
	The role of the SAC with respect to this dimension is hard to determine (not apply to the catalogues) or its unknown.

3.2.3. FAIR-enabling analysis methodology

Our FAIR-enabling assessment voluntarily stayed very coarse. Our motivation was not to compare the SACs but more to show which dimensions are addressed by which type of SACs, where SAC may have to put efforts in the future, and finally, emphasise specific aspects of certain SAC so they could serve as an example. To assess the FAIR-enabling level as well as we can, we apply multiple strategies:

- The following repositories were assessed by their mainteners or developers who are member of FAIR-IMPACT's WP4 and the assessment were discussed during task meetings:
  - SIFR BioPortal (INRAE)
  - AgroPortal (INRAE)
  - EcoPortal (LifeWatch)
  - EarthPortal (Data Terra)
  - LOV (UPM)
  - Identifiers.org (UNIMAN)
  - OEG Vocabularies (UPM)

Funded by the European Union

    - (to come) IVOA Vocabularies (OBS-PARIS)
    - (to come) CESSDA Vocabulary Service
    - (to come) Vocabs Data Stations
- We ignored the prototypes and retired SACs.
- All the SACs of type “Listing” and “Search service” where assessed by their type i.e., it is the fact of being such type of SAC that induces the FAIR-enabling capabilities. Typically, listings do not support any of the 10 FAIR-enabling dimensions. Search services only help with the Cataloguing dimension.
- Many SACs were grouped by technology and assessed based on our task members knowledge of the technology. This applied to SAC based on the following technologies: OLS, ORR, ShowVoc, VocPrez, and to some extent OntoPortal (except the ones not listed above).
- Each time we were aware of a particular enforcement or attention a SAC would pay a specific attention to, we changed the assessment that was based on the technology. For instance, we know EcoPortal rigorously checks the licensing info of its resources and makes it a requirement to get into EcoPortal, so the criteria went from orange (as other OntoPortals) to green in EcoPortal’s case.

Over the 112 “Active” SAC in our listing, we assessed 59 of them. We focused on the most known and prominent catalogues and did not target to assess them all. In the future, we will leave a new version of the associated spreadsheet open for comment, so we could take contributions and comments from any ones outside FAIR-IMPACT, being a maintainer or not of a catalogue to improve this work both by: (i) completing the listing of SACs and (ii) refining the FAIR-enabling assessment.

Funded by the European Union

4. Results

4.1. Comprehensive listing of Semantic Artefact Catalogues

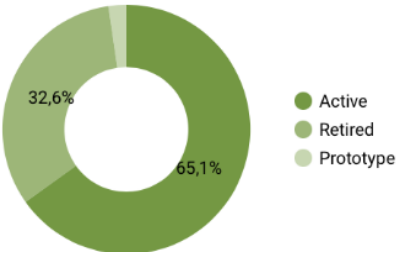
After filtering by status, the following section presents an analysis of all SACs with an "Active" status related to their types, disciplines and technologies. Further, we provide a cross analysis between the disciplines of the active catalogues and either their types or the generic technology used. For this analysis, we used the tool *Looker Studio* which allowed us to produce customizable, informative reports and dashboards from GSheets.
The complete list of all the SAC gathered and analysed corresponds to the version 1 of the associated spreadsheet data, doi:10.5281/zenodo.12799862. **We were able to review a total of 172 semantic artefact catalogues.**

4.1.1. Status

Over 112 "Active", 56 "Retired", and 4 "Prototype" catalogues were reviewed (Fig. 1). This indicates that a significant proportion (65.1%) of the gathered SACs are currently usable and this figure (112) demonstrates the importance of their missions. Among the 32% of the "Retired" SACs some can still be used and are still accessible (but often outdated) and others are not accessible at all anymore.

Figure 1. Number and distribution of all SAC listed by status.

	Status	Record Count ▼
1.	Active	112
2.	Retired	56
3.	Prototype	4



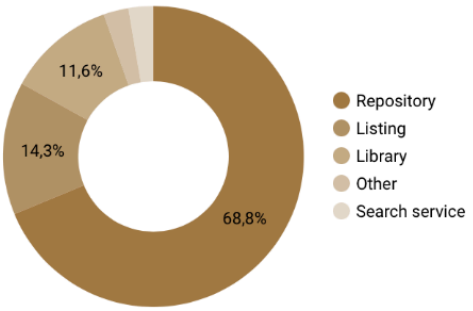
4.1.2. Types of SAC

Among the 112 "Active" SAC" collected, the "Repository" type is the most prevalent, representing nearly 70% of the reviewed SACs (Fig. 2). The "Listing" and "Library" types are similarly represented in our listing, comprising 14% and 11% of the SAC, respectively. Finally, the "Search Service" type is the least represented, with only 3 SACs. Finally, 3 catalogues were classified as "Other" as we believed they were slightly different from the enumerated types, but still deserved to be mentioned in our study. They are: OntologyDesignPatterns.org –a portal dedicated to ontology design patterns (ODPs), Prefix.cc –a namespace lookup service, and Identifiers.org –a tool to resolve compact identifiers. The number of repositories (77) shows the importance of having concrete applications serving the content of SA and not only their metadata.

Figure 2. Number and distribution of all SAC listed with active status by types.

Funded by the European Union

	Type	Record Count
1.	Repository	77
2.	Listing	16
3.	Library	13
4.	Other	3
5.	Search service	3



4.1.3. Generic technology

Regarding the technology used by the SACs, we identified 15 different technologies (Table 2) but found instances of 12 of these technologies (Fig. 3). This is due to the fact that 3 private technologies (Centree, Data Harmony Taxonomy Suite and OCLC Terminology Services) did not make public the use of their technologies (usually their clients).[32] For SACs that were unable to determine the technology (and its genericity), we categorise the undefined ones in the “Other” category" which represents more than half (55%) of the SACs.

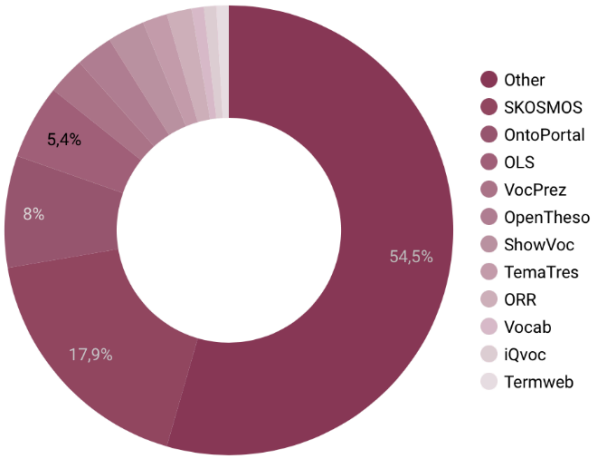
The top 3 most prevalent known technologies in our list are SKOSMOS, OntoPortal and OLS, which represent respectively 17.9%, 8% and 5.4% of the SAC listed. Despite being limited to a certain type of semantic artefacts (i.e., the ones represented in SKOS), SKOSMOS is known to be the technology the most focused and easy to install, which indirectly explains the fact that it's the most reused. Often, the catalogues built with SKOSMOS are set up by an infrastructure or project and are not targeting to be “reference SAC'' for a discipline. Typically, all the SKOMOS based SACs never accept direct SA submission from external parties (not running the SAC). We may also observe that the OntoPortal technology, for which the OntoPortal Alliance works together to develop a shared open source technology for SACs, is well adopted too despite the fact that it is known to be complex. Most of the cases of the use of OntoPortal are to develop reference discipline open SACs where SA submissions can be accepted from anyone. We may also note that 4 of the reuses (over 5) of the OLS technology are done by one big infrastructure program (N4DI via TIB) showing less diversity of reuses than OntoPortal.

Figure 3. Number and distribution of all SAC listed with active status by generic technology used.

[32] Note that we have identified these 3 private technologies but there could be more. This is beyond the scope of this work to review private solutions.

Funded by the European Union

	Generic technology	Record Count
1.	Other	61
2.	SKOSMOS	20
3.	OntoPortal	9
4.	OLS	6
5.	VocPrez	3
6.	OpenTheso	3
7.	ShowVoc	3
8.	TemaTres	2
9.	ORR	2
10.	Vocab	1
11.	iQvoc	1
12.	Termweb	1



4.1.4. Disciplines

We identified 17 disciplines and the "General" category for the list of catalogues and this number remains stable when considering only the "Active" SAC. The "General" category accounts for 40% of SACs (Fig. 4). This high proportion is primarily due to the broad range of applications and usages of SA covered by this category. It basically translates that 40% of the SACs reviewed are not domain specific thus are targeting users from various communities or are listing/hosting SAs that can be used in multiple disciplines (e.g., general metadata vocabularies, standards, etc.)

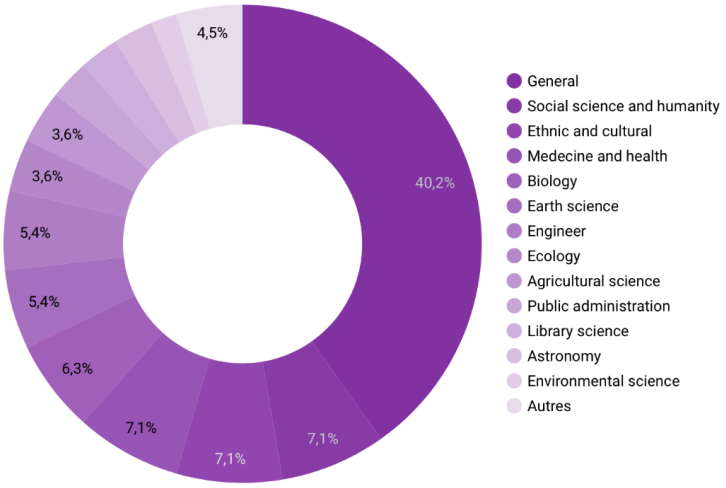
The most represented disciplines are "Social science and humanities", "Ethnic and cultural sciences", "Medicine and health" where each represent 7% of the reviewed SACs' disciplines, followed by "Biology", "Earth science" and "Engineer" each account for 6.3%, 5.4% and 5.4% respectively. Together, these six disciplines comprise just more than a third (38.4%) of all reviewed actives SACs. A second gap is observed among disciplines that represent 4 or less SAC. This final quarter (21.4%) includes more than half (11 out of 17) of the disciplines identified during our study. Within this segment, we observe two distinct groupings: The first includes "Ecology", "Agricultural science", "Library Science", "Library science", "Public administration" and "Astronomy" each accounting for 3% or less of the SACs. The second includes six disciplines –"Linguistic", "Economics", "Material science", "Education" and "Chemistry"– where there is a unique SAC.

Although "Ecology" is in Wikipedia a sub-discipline of "Biology", we chose to list it separately due to its significance. If "Ecology" were included within the "Biology" field, "Biology" would become the second most represented domain with 11 SAC. Another observation is the prominence of the "Medicine and health" and "Biology" disciplines that together make 15 of the reviewed actives SACs; indeed biomedicine has always been very active and productive in ontologies and the semantic Web as illustrated by historical and important catalogues such as BioPortal, OLS, OntoBee, AberOWL, OBO Foundry, etc.

Figure 4. Number and distribution of all SAC listed with active status by discipline domains.

Funded by the European Union

	Standardized Wikipedia Discipline	Record Count
1.	General	45
2.	Medecine and health	8
3.	Ethnic and cultural	8
4.	Social science and humanity	8
5.	Biology	7
6.	Earth science	6
7.	Engineer	6
8.	Ecology	4
9.	Agricultural science	4
10.	Library science	3
11.	Public administration	3
12.	Astronomy	3
13.	Environmental science	2
14.	Linguistic	1
15.	Economics	1
16.	Material sciences	1
17.	Education	1
18.	Chemistry	1



4.1.5. Cross-analysis of disciplines by type and by technology

The following subsections presents a cross-analysis of the disciplines of the active reviewed SACs, classified by type and by technology. In both subsections, the "General" category has to be considered specifically.

By type

Here, the aim is to examine the distribution of various types of SACs across different disciplines. We observe a bipartite repartition of the technology used by the disciplines. The first part of the disciplines are covered by only one type of SAC (7 out of 17) while the second part of disciplines are covered by two types of SAC (8 out of 17), with the "Repository" type being the most prevalent (Fig. 5).

The "Repository" type is adopted by "Agricultural Science", "Chemistry", "Earth Science", "Economics", “Linguistic”, "Material Science" and "Public Administration". However, "Education" is the unique discipline that utilises a single SAC type which is not “Repository” but the "Listing" type. "Ecology", and "Medicine and health" disciplines use both the "Repository" and "Listing" types, while "Biology", “Engineer”, “Library science” and "Social science and humanities" employ the "Repository" and "Library" types. However "Ethnic and Cultural Science" is the only one that encompasses "Repository" and "Search Service" types and "Environmental Science" is also unique in that it uses two types of SAC: "Listing" and "Library", without including the "Repository" type. Finally, there is only one discipline - “Astronomy” - that uses three types of SAC, which are “Repository”, “Listing” and “Library”.

Figure 5. Cross-analysis of disciplines of active SACs by types.

Funded by the European Union

				Type / Record Count	
Standardized Wikipedia Discipline	Search service	Repository	Other	Listing	Library
Agricultural science	-	4	-	-	-
Astronomy	-	1	-	1	1
Biology	-	6	-	-	1
Chemistry	-	1	-	-	-
Earth science	-	6	-	-	-
Ecology	-	3	-	1	-
Economics	-	1	-	-	-
Education	-	-	-	1	-
Engineer	-	4	-	-	2
Environmental science	-	-	-	1	1
Ethnic and cultural	1	7	-	-	-
General	2	23	3	11	6
Library science	-	2	-	-	1
Linguistic	-	1	-	-	-
Material sciences	-	1	-	-	-
Medecine and health	-	7	-	1	-
Public administration	-	3	-	-	-
Social science and humanity	-	7	-	-	1

Overall, this cross-analysis shows us that all disciplines are covered by one or several repositories (except “Education”) and some disciplines may even need to converge their efforts as multiple repositories (too much?) are available in their area at the risk of maybe confusing final users: “Biology”, “Earth Sciences”, “Ethnic and cultural”, “Medicine and health” and “Social science and humanity”.

By technology

Here, the aim is to examine the distribution of various generic technologies of SAC across different disciplines. We exclude the "Other" category as it does not pertain to specific technology but rather to a group of undefined or prototype technologies, even though it represents more than 50% (Fig. 3) of the SACs. Furthermore, we exclude "Astronomy", "Education" and "Environmental Science" from this analysis, since the technologies used for their SACs fall into the "Other" category (Fig. 6). Consequently, 14 out of the 17 disciplines are analysed. However, disciplines that predominantly utilised a type of technology categorised as “Other,” like “Engineering,” “Ethnic and Cultural Studies,” and “Medicine and Health,” are analysed but with moderate confidence.

As seen before (Fig. 3), SKOSMOS is the most prevalent technology; here (Fig. 6) we can see that with the exception of the “Social science and humanity” discipline, SKOSMOS is used mostly once, when used, in other disciplines. This may be interpreted as a call for the “Social science and humanity” to harmonise and converge the semantic artefact catalogues used in this area. This sur-representation of a technology within a certain discipline is not found with the OntoPortal and OLS that were the two other well adopted technologies.

Fig. 6 also shows that "Earth Science" is the only discipline using four different types of technology: SKOSMOS, OntoPortal, VocPrez, and ORR; which seems to also call for an harmonisation of the semantic artefact catalogues used in this area. To some extent, this is also a bit true for “Engineer" which uses three types of technologies. Five out of the

Funded by the European Union

fourteen domains use two types of technology. We can see some technologies are used only in some disciplines e.g., OpenTheso being only used in the Social science and humanity”” and “Ethnic and cultural” or VocPrez and ORR being only used in the “Earth science” disciplines. Finally, 6 out of the 14 disciplines use a unique technology. “Economic”, "Linguistics" and “Public administration” use only SKOSMOS, "Library science" uses iQvo, "Chemistry" uses OLS, and "Material Science" uses OntoPortal.

Figure 6. Cross-analysis of disciplines of active SACs by technology.

	Generic technology / Record Count											
Standardized Wikipedia Dis…	iQvoc	Vocab	VocPrez	Termweb	TemaTres	ShowVoc	SKOSMOS	Other	OpenThe…	OntoPortal	ORR	OLS
Agricultural science	-	-	-	-	-	1	-	2	-	1	-	-
Astronomy	-	-	-	-	-	-	-	3	-	-	-	-
Biology	-	-	-	-	-	-	-	3	-	2	-	2
Chemistry	-	-	-	-	-	-	-	-	-	-	-	1
Earth science	-	-	2	-	-	-	1	-	-	1	2	-
Ecology	-	-	-	-	-	1	-	1	-	2	-	-
Economics	-	-	-	-	-	-	1	-	-	-	-	-
Education	-	-	-	-	-	-	-	1	-	-	-	-
Engineer	-	-	-	-	-	-	1	3	-	-	-	1
Engineer	-	-	-	-	-	-	-	-	-	1	-	-
Environmental science	-	-	-	-	-	-	-	2	-	-	-	-
Ethnic and cultural	-	-	-	-	1	-	-	6	1	-	-	-
General	-	1	1	1	1	1	9	30	-	-	-	1
Library science	1	-	-	-	-	-	-	2	-	-	-	-
Linguistic	-	-	-	-	-	-	1	-	-	-	-	-
Material sciences	-	-	-	-	-	-	-	-	-	1	-	-
Medecine and health	-	-	-	-	-	-	-	6	-	1	-	1
Public administration	-	-	-	-	-	-	2	1	-	-	-	-
Social science and humanity	-	-	-	-	-	-	5	1	2	-	-	-

4.2. FAIR-enabling dimension

Here, we analyse a few lessons learnt from the assessment of the FAIR-enabling dimensions for 59 over 112 (53%) active SACs in our listing. Especially, the FAIR-enabling dimensions assessment per type of SACs is synthesised in Fig. 7.

4.2.1. Type of SAC

None of the FAIR-enabling dimensions are addressed by the “Listing” type of SAC. All dimensions were assessed to level 3 (Table 3). Indeed, the fact of listing semantic artefacts on a web page somewhere does not concretely help or support to make an artefact FAIR and we found that most of the criterias within a FAIR-enabling dimension were not fulfilled by “Listings”.

“Search service” and “Other” types of SAC do not support FAIR either with the exception of the criteria “Cataloguing” that was assessed to level 2 for most of them. Indeed, this acknowledges that when a semantic artefact is searchable by these SACs, it eases their finability. Still, we have not identified any active “Search service” that would completely take care of this cataloguing dimension (level 1) as it used to be done by retired semantic web search engines such as Swoogle and Watson.

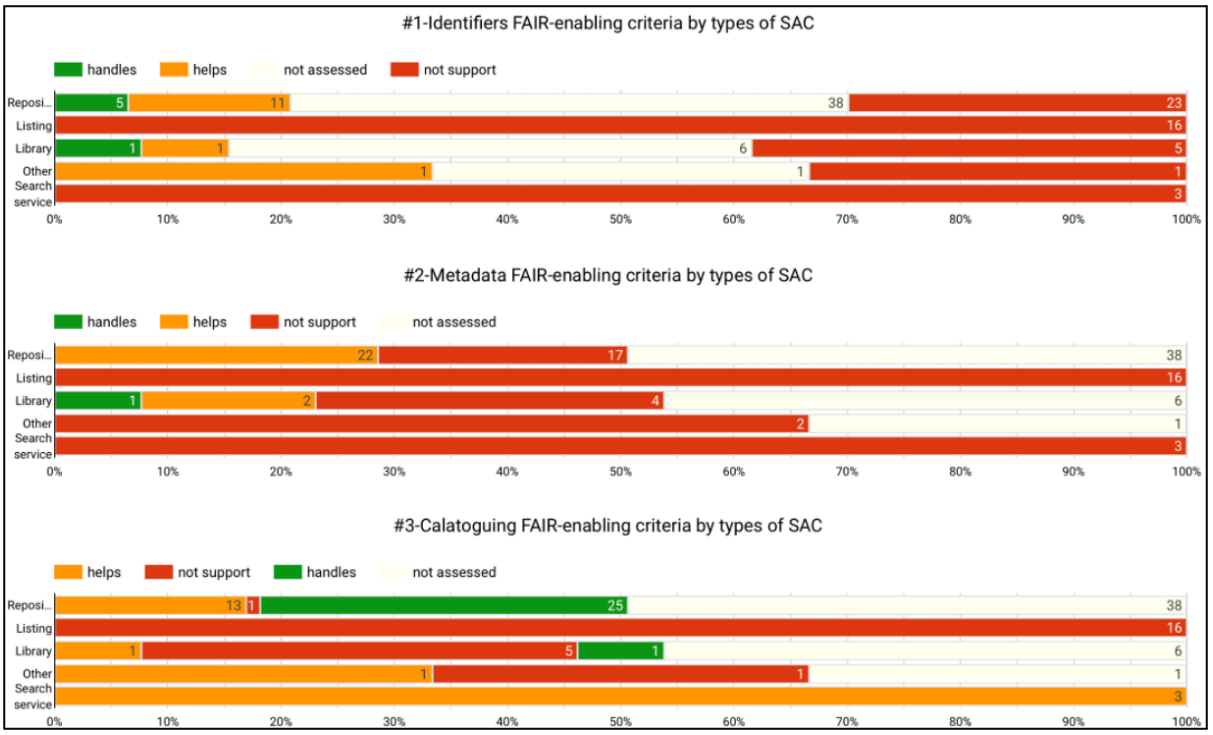
Over the 7 over 13 “Library” type of SAC assessed, we found that most of them do not support most FAIR-enabling dimensions to the exception of the OBO Foundry and OEG

Funded by the European Union

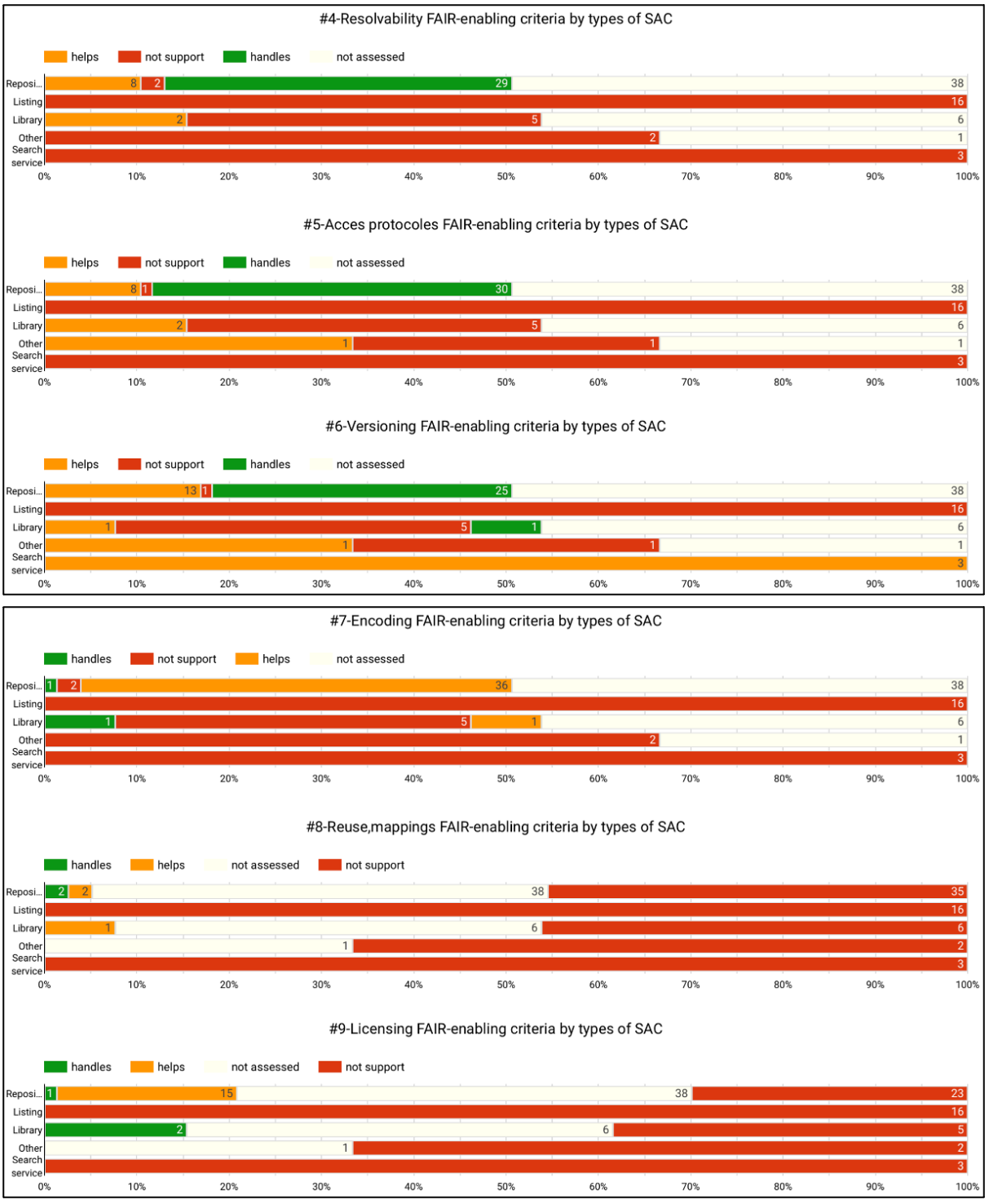
Vocabularies. These 2 SACs are good counter examples of "Library" either maintained respectively by a wide, well organised community or by a lab team; in both cases the mainteners have set up the mechanisms to take care (level 1) of some multiple FAIR-enabling dimensions. For example, when an ontology is incorporated in the OBO Foundry, after a peer-review quality check, the identifiers, metadata, cataloguing, or licensing aspects are highly facilitated or managed by the Foundry and ontology developers only have to follow the procedures.

Finally, we have assessed 39 over 77 (51%) of the SAC of type **"Repository". In this type of SAC, multiple FAIR-enabling dimensions are diversely addressed**. Typically, the 3 dimensions Cataloguing, Resolvability, and Access Protocols are well generally well addressed (often at level 1) by SACs, which sounds logical are the 3 dimensions correspond to the primary roles of SACs of cataloguing semantic artefacts, indexing and searching their content and making their content accessible with standard protocols. Sometimes, a SAC will put an emphasis on a dimension making it level 1 where most of the SAC are level 2. For instance, Identifiers in the NERC Vocabulary Service, Licensing in EcoPortal, Versioning in AgroPortal or Reuse and mappings in LOV or Planteome. Overall, we could say SAC of type "Repository" really help/support, at least up to level 2, multiple FAIR-enabling dimensions; still some dimensions are mostly poorly addressed (level 3) such as Versioning, Reuse and mappings and Provenance. This analysis shows SAC mainteners where to put some energy now in order to continue and enhance their support to semantic artefact developers in making their resources FAIR.

Figure 7. FAIR-enabling dimensions assessment per type of SACs



Funded by the European Union



4.2.2. Generic technology

Some technologies support more or less FAIR-enabling dimensions. For instance, the 3 dimensions Cataloguing, Resolvability, and Access protocols are naturally well supported by “OntoPortal” and “OLS” or “VocPrez”. The only technology where the Versioning dimension is always assessed level 1 or 2 is for “OntoPortal” which shows that this technology is the

Funded by the European Union

only one dealing with SA versioning criterias. Catalogues such as AgroPortal, EcoPortal and EarthPortal (all involved in FAIR-IMPACT) even explicitly declared a level 1 support for Versioning.
Similarly, some dimensions are not addressed at all by any of the SACs running a given technology e.g., Reuse and mappings. The Encoding dimension is always assessed to level 2 for SAC of type repository, whatever their technology: this means all SAC of type repository consider as a requirement the fact of adopting semantic web standards (OWL, SKOS) to encode SA and thus being hosted/served by the SAC.
The rest of the dimensions (Identifiers, Metadata, Licensing and Provenance) vary independently of the technology used. This shows that even when relying on the same technology, SAC can decide to enforce (as an editorial guideline or community enforcement, or technological complement) certain dimensions more than others. This is specifically true for “SKOSMOS” where certain installations (e.g. Loterre) emphasise Identifiers or Metadata dimensions and this is independent of the use of the “SKOSMOS” technology.

Funded by
the European Union

5. Conclusions and next steps

In the rapidly evolving landscape of scientific research, the effective management and utilisation of ontologies and other semantic artefacts are crucial. Semantic Artefact Catalogues (SACs) are indispensable tools in this regard, particularly within the framework of the EOSC. These catalogues provide essential platforms for receiving, hosting, serving, aligning, and enabling the reuse of SAs. By facilitating the organisation and access of these artefacts, SACs support and often ensure their compliance with the FAIR Principles, which are foundational to the EOSC's mission of promoting open science and data sharing across diverse scientific disciplines.

Within the FAIR-IMPACT project, particularly WP4 on ontologies and metadata, we focused on establishing guidelines and community practices for the lifecycle of FAIR semantic artefacts. This included their creation, sharing, reuse, and the standardisation of SA metadata descriptions and SAC application programming interfaces. Our collaborative efforts with multiple communities, such as AgroPortal, EcoPortal, and EarthPortal, demonstrated the practical application of these guidelines and tools, promoting the interoperability and FAIRness of semantic artefacts across different domains.

This report describes the process followed to provide an exhaustive listing of SAC and their various features and the methodology used to define FAIR-enabling dimensions. Our comprehensive review of SACs, encompassing current and past catalogues, highlights the critical role these resources play in the FAIR ecosystem. We sorted SACs by types ((from simple listings to advanced repositories), disciplines, and technologies, and identified 10 key dimensions for assessing their FAIR-enabling capabilities using methodologies and tools such as O'FAIRe, FOOPS!, FsF, 10-SR, and FVF. This analysis demonstrates how SACs facilitate the adherence of semantic artefacts to FAIR principles. The findings presented in this milestone provide a detailed overview of the current landscape of SACs within the EOSC and beyond. This diversity reflects the tailored approaches different communities have adopted to meet their specific needs.

The milestone comprises the current report, which details our methodology and analysis results, along with an associated spreadsheet. The spreadsheet includes a list of SACs, their classifications (status, type, discipline, technology), and the evaluation of their FAIR-enabling dimensions. **The version of the spreadsheet discussed in this report is accessible with the DOI: 10.5281/zenodo.12799862 and a link to the current "live" spreadsheet is provided with the description of this frozen spreadsheet (v1) for future comments and suggestions so we could produced completed and corrected version of these data**.

In conclusion, our work underscores the necessity of robust SACs for the efficient management and utilisation of semantic artefacts, crucial for advancing the EOSC's vision of a Web of FAIR Data and Services. By providing a detailed assessment framework and comprehensive review, we aim to facilitate the selection and use of SACs, promoting the harmonisation and synchronisation of FAIR-enabling practices. This milestone represents a significant step towards achieving a more interconnected and open scientific data ecosystem, fostering collaboration and innovation across diverse research disciplines.

Funded by the European Union

6. Appendices

6.1. Grouping of the criterias, questions of recommendations

Here we regroup the criterias, questions or recommendations of the 5 state-of-the-art tools and methods. When a line is bolded it recognizes that the corresponding criteria, question or recommendation is directly addressed by the catalogue and does not depend directly on the SA itself (e.g., its content or metadata).

6.1.1. Identifiers

- F1Q1 Does the ontology have a "local" identifier, i.e., a globally unique and potentially permanent identifier assigned by the developer (or developing organization)?
- F1Q2 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?
- PURL1: This check verifies if the ontology has a persistent URL (w3id, purl, DOI, or a W3C URL)
- URI2: This check verifies if the ontology URI is equal to the ontology ID
- ****F1Q3 Are the ontology metadata clearly identified either by the same identifier than the ontology (if included in the ontology file) or with its own globally unique and persistent identifier?**
- F1Q4 Does the ontology provide a version-specific URI, and is this URI resolvable?
- VER1: This check verifies if there is an id for this ontology version, and whether the id is unique (i.e., different from the ontology URI)
- P-Rec. 1Globally Unique, Persistent and ResolvableIdentifiers must be used for Semantic Artefacts, their content (terms/ concepts/ classes and relations), and their version.
- P-Rec. 2Globally Unique, Persistent, andResolvable Identifiers must be used for SemanticArtefact Metadata Records. Metadata and data must be published separately, even if it is managed jointly.
- Rule 5. Assign a unique and persistent identifier to (a) the vocabulary and
- (b) each term in the vocabulary
- FVF-1 Vocabulary and constituent terms are assigned globally unique and persistent identifiers.

6.1.2. Metadata

- F2Q1 Is the ontology described with additional 'MIRO must' metadata properties?
- OM1: This check verifies if the The following minimum metadata [title, description, license, version iri, creator, creationDate, namespace URI] are present in the ontology
- F2Q2 "Is the ontology described with additional 'MIRO should' or 'optional' metadata properties?
- F2Q3 Is the ontology described with another metadata property with no explicit corresponding MIRO requirement?
- FIND1: This check verifies if an ontology prefix is available

Funded by the European Union

- OM2: This check verifies if the following recommended metadata [NS Prefix, version info, creation date, citation] are present in the ontology.
- OM3: This check verifies if the following detailed metadata [doi, publisher, logo, status, source, issued date] are present in the ontology.
- ****F3Q1 Are the ontology metadata included and maintained in the ontology file?**
- ****F3Q2 If not, are the ontology metadata described in an external file?**
- ****F3Q3 Does that external file explicitly link to the ontology and vice-versa?**
- P-Rec. 3A common minimum metadata schema must be used to describe semantic artefacts and their content.
- Rule 7. Add vocabulary metadata
- FVF-2 Vocabulary and constituent terms have rich metadata.
- R1Q1 Does the ontology provide information about how classes or concepts are defined?
- R1Q2 Does the ontology provide metadata information about its hierarchy?
- R1Q3 How much of the ontology objects are described with labels?
- VOC3: This check verifies the extent to which all ontology terms have labels (rdfs:label in OWL vocabularies, skos:prefLabel in SKOS vocabularies)
- R1Q4 How much of the ontology objects are defined using a text description?
- VOC4: This check verifies whether all ontology terms have descriptions (rdfs:comment in OWL vocabularies, skos:definition in SKOS vocabularies)
- R1Q5 How much ontology objects are defined using a property restriction or an equivalent class?
- R1Q6 How much ontology objects provide provenance information with annotation properties (e.g., author, date)?
- Rule 3. Check term and definition completeness and consistency in the
- legacy vocabulary
- FVF-9 Vocabulary and constituent terms are described with a plurality of accurate and relevant attributes.

6.1.3. Cataloguing

- F4Q1 Is the ontology registered in multiple ontology 'libraries'?
- F4Q2 Is the ontology registered in multiple open ontology 'repositories'?
- FIND3: This check verifies if the ontology can be found in a public registry (LOV)
- FIND2: This check verifies if the ontology prefix can be found in prefix.cc or LOV registries. This check also verifies if the prefix resolves to the same namespaceprefix found in the ontology.
- ****P-Rec. 4 Semantic Artefact and its content should be published in a trustworthy semantic repository.**
- ****F4Q3 Are the ontology 'libraries' or 'repositories' properly indexed by Web search engines?**
- ****P- Rec. 6 Build semantic artefact search engines that operate across different semantic repositories.**
- Rule 4. Establish a traceable maintenance-environment for the FAIR

Funded by the European Union

- vocabulary content
- Rule 8. Register the vocabulary
- FVF-4 Vocabulary and constituent terms are registered or indexed in a searchable engine or a resource.

6.1.4. Resolvability

- A1Q1 Do the ontology URI and other identifiers, if they exist, resolve to the ontology?
- URI1: This check verifies if the ontology URI found within the ontology document is resolvable
- VER2: This check verifies if the version IRI resolves
- ****A1Q2 Does the ontology URI (if metadata are included in the ontology file) or the external metadata URI resolve to the metadata record?**
- A1Q3 Do the ontology URI and the external metadata URI (if the metadata are not included in the ontology file), support content negotiation?
- CN1: This check verifies of the ontology URI is published following the right content negotiation for RDF and HTML
- RDF1: This check verifies if the ontology has an RDF serialization (ttl, n3, rdf/xml, json-ld)
- ****P-Rec. 5 Semantic repositories must offer access to Semantic Artefacts and their content using community standard APIs and serializations to support both use/ reuse and indexation by search engines.**
- P-Rec. 9Semantic artefacts must be made available as a minimum portfolio of common serialization formats.
- Rule 6. Create machine readable representations of the vocabulary terms
- Rule 9. Make the vocabulary accessible for humans and machines
- FVF-3 Vocabulary and constituent terms can be accessed using identifiers, preferably by both humans and machines.

6.1.5. Access protocols

- A1Q4 Are the ontology and its metadata accessible through another standard protocol such as SPARQL?
- ****A1.1Q1 Is the ontology relying on HTTP/URIs for its identification and access mechanisms?**
- ****A1.1Q2 Is the ontology access protocol open, free, and universally implementable?**
- HTTP1: This check verifies if the ontology uses an open protocol (HTTP or HTTPS)
- ****A1.1Q3 If the ontology and metadata are accessible through another protocol, is that protocol open, free, and universally implementable?**
- ****A1.2Q1 Is the ontology accessible through a protocol that supports authentication and authorization?**
- ****A1.2Q2 Are the ontology metadata accessible through a protocol that supports authentication and authorization?**
- ****P-Rec. 7 Repositories must offer a secure access protocol, and appropriate user access control functionalities.**

Funded by the European Union

- FVF-5 Vocabulary and constituent terms are retrievable using a standardised communication protocol, preferably open, free and universally implementable protocols, which allow for authentication and authorisation, where necessary.

6.1.6. Versioning

- ****A2Q1 Is the ontology accessible in a repository that supports versioning?**
- ****A2Q2 Are the ontology metadata of each version available?**
- ****A2Q3 Are the ontology metadata accessible even if no more versions of the** ontology are available?
- A2Q4 Is the status of the ontology clearly informed?
- FIND_3_BIS: Metadata are accessible even when the ontology is no longer available. Since the metadata is usually included in the ontology, this check verifies whether the ontology is registered in a public metadata registry (LOV)
- P-Rec. 8Human and machine-readable persistence policies for semantic artefacts metadata and data must be published.
- Rule 10. Implement a process for publishing revisions of the FAIR
- vocabulary
- FVF-6 Vocabulary and constituent terms are persistent over time and are appropriately versioned.

6.1.7. Encoding

- I1Q1 What is the representation language used for the ontology and ontology metadata?
- I1Q2 Is the representation language used in a W3C Recommendation?
- P-Rec. 11 A standardized knowledge representation language should be used for describing semantic artefacts.
- I1Q3 Is the syntax of the ontology informed?
- I1Q4 Is the formality level of the ontology informed?
- I1Q5 Is the availability of other syntaxes/formats informed?
- FVF-7 Vocabulary and constituent terms use a formal, accessible and broadly applicable, and preferably machine-understandable language for knowledge representation.

6.1.8. Reuse, mappings & crossrefs

- I2Q1 Does the ontology import other FAIR vocabularies?
- I2Q2 Does the ontology reuse terms from other FAIR vocabularies (URIs)?
- VOC2: This check verifies if the ontology imports/extends other vocabularies (besides RDF, OWL and RDFS)
- I2Q3 If yes, does it include the minimum information for those terms?
- FVF-8 Vocabulary and constituent terms use qualified references to other vocabularies.
- I2Q4 Is the ontology aligned to other FAIR vocabularies?
- I2Q5 If yes, are those alignments well represented and to unambiguous entities? If yes, are those alignments curated?
- P-Rec. 10 Foundational Ontologies may be used to align semantic artefacts.

Funded by the European Union

- P-Rec. 12 Semantic mappings between the different elements of semantic artefacts should be published in machine-readable formats.
- I2Q6 Does the ontology provide information about the relation to or influence of other FAIR vocabularies?
- P-Rec. 13 Crosswalks, mappings and bridging between semantic artefacts should be documented, published and curated.
- I3Q1 Does the ontology provide qualified cross-references to external resources/databases?
- I3Q2 If yes, are those cross-references well represented and to unambiguous entities?
- I3Q3 Does the ontology use valid URIs to encode some metadata values?
- I2Q7 Does the ontology reuse standard and FAIR metadata vocabularies to describe its metadata?
- VOC1: This check verifies if the ontology reuses other vocabularies for declaring metadata terms
- P-Rec. 14 Standard vocabularies should be used to describe semantic artefacts.

6.1.9. Licensing

- R1.1Q1 Is the ontology license clearly specified, with an URI that is resolvable and supports content negotiation?
- OM4.1: This check verifies if a license associated with the ontology
- OM4.2: This check verifies if the ontology license is resolvable
- P-Rec. 16 The semantic artefact must be clearly licenced for use by machines and humans.
- R1.1Q2 Are the ontology access rights specified and permissions documented?
- R1.1Q3 Are the ontology usage guidelines and copyright holder documented?
- Rule 2. Verify that the legacy-vocabulary license allows repurposing, and
- agree on the license for the FAIR vocabulary
- FVF-10 Vocabularies are released with a standard data usage licence, preferably a machine-readable licence.
- Rule 1. Determine the governance arrangements and custodian of the
- legacy vocabulary
- ****R1.3Q3 Is the ontology openly and freely available?**

6.1.10. Provenance

- R1.2Q1 Does the ontology provide information about the actors involved in its development?
- R1.2Q2 Does the ontology provide information about its general provenance?
- OM5_1: This check verifies if basic provenance is available for the ontology: [author, creation date].
- OM5_2: This check verifies if detailed provenance information is available for the ontology: [issued date, publisher

Funded by the European Union

- P-Rec. 15 Provenance information regarding the reuse of components from third-party semantic artefacts should be made explicit.
- P-Rec. 17 Provenance must be clear for both humans and machines.
- R1.2Q3 Are the accrual methods and policy of the ontology documented?
- R1.2Q4 Is the ontology clearly versioned with version information and links to previous versions?
- R1.2Q5 Are the ontology latest changes documented?
- R1.2Q6 Are the methodology and tools used to build the ontology documented?
- R1.2Q7 Is the ontology rationale documented?
- DOC1: This check verifies if the ontology has an HTML documentation
- R1.2Q8 Does the ontology inform about its funding organization?
- R1.3Q1 Does the ontology provide information about projects using or organizations endorsing?
- ****R1.3Q2 Is the ontology included in a specific community set or group?**
- FVF-11 Vocabularies meet domain-relevant community standards.

7. References in Section 2 (from ISWC 2023 paper)

1. Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C. 't, Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Barend Mons: The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 3, (2016). https://doi.org/10.1038/sdata.2016.18.
2. Poveda-Villalón, M., Espinoza-Arias, P., Garijo, D., Corcho, O.: Coming to Terms with FAIR Ontologies. In: Keet, M.C. and Dumontier, M. (eds.) 22nd International Conference on Knowledge Engineering and Knowledge Management, EKAW'20. pp. 255–270. Springer, Bolzano, Italy (2020). https://doi.org/10.1007/978-3-030-61244-3_18.
3. Amdouni, E., Jonquet, C.: FAIR or FAIRer? An integrated quantitative FAIRness assessment grid for semantic resources and ontologies. In: Emmanouel Garoufallou and María-Antonia Ovalle-PerandonesAndreas Vlachidis (eds.) 15th International Conference on Metadata and Semantics Research, MTSR'21. pp. 67–80. Springer, Madrid, Spain (2021). https://doi.org/10.1007/978-3-030-98876-0_6.
4. Coxid, S.J.D., Gonzalez-Beltranid, A.N., Magagna, B., Marinescu, M.-C.: Ten simple rules for making a vocabulary FAIR. (2021). https://doi.org/10.1371/journal.pcbi.1009041.
5. Jonquet, C., Toulet, A., Dutta, B., Emonet, V.: Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal. Data Semantics. 7, 191–221 (2018). https://doi.org/10.1007/s13740-018-0091-5.
6. Amdouni, E., Bouazzouni, S., Jonquet, C., O'faire, C.J.: O'FAIRe makes you an offer: Metadata-based Automatic FAIRness Assessment for Ontologies and Semantic Resources. Int J Metadata Semant Ontol. 16, 16–46 (2022). https://doi.org/10.13039/501100001665.
7. Graybeal, J., Jonquet, C., Fiore, N., Musen, M.A.: Adoption of BioPortal's Ontology Registry Software: The Emerging OntoPortal Community. In: 13th Plenary Meeting of the Research Data Alliance, Poster session. , Philadelphia, PA, USA (2019).

Funded by the European Union

8. Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N.B., Jonquet, C., Rubin, D.L., Storey, M.-A., Chute, C.G., Musen, M.A.: BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res. 37, 170–173 (2009). https://doi.org/10.1093/nar/gkp440.
9. Whetzel, P.L., Noy, N.F., Shah, N.H., Alexander, P.R., Nyulas, C., Tudorache, T., Musen, M.A.: BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Res. 39, 541–545 (2011). https://doi.org/10.1093/nar/gkr469.
10. Kechagioglou, X., Vaira, L., Tomassino, P., Fiore, N., Basset, A., Rosati, I.: EcoPortal: An Environment for FAIR Semantic Resources in the Ecological Domain. In: 3rd International Workshop on Semantics for Biodiversity (S4BioDiv'2021). CEUR, Bolzano, Italy (2021).
11. Ding, Y., Fensel, D.: Ontology Library Systems: The key to successful Ontology Re-use. In: 1st Semantic Web Working Symposium, SWWS'01. pp. 93–112. CEUR-WS.org, Stanford, CA, USA (2001).
12. Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J., Consortium, T.O.B.I., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R.H., Shah, N.H., Whetzel, P.L., Lewis, S.: The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 25, 1251–1255 (2007). https://doi.org/10.1038/nbt1346.
13. Hartmann, J., Palma, R., Gómez-Pérez, A.: Ontology Repositories. Handbook on Ontologies. 551–571 (2009). https://doi.org/10.1007/978-3-540-92673-3.
14. Ong, E., Xiang, Z., Zhao, B., Liu, Y., Lin, Y., Zheng, J., Mungall, C., Courtot, M., Ruttenberg, A., He, Y.: Ontobee: A linked ontology data server to support ontology term dereferencing, linkage, query and integration. Nucleic Acids Res. 45, D347–D352 (2016). https://doi.org/10.1093/nar/gkw918.
15. Côté, R.G., Jones, P., Apweiler, R., Hermjakob, H.: The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries. BMC Bioinformatics. 7, 97 (2006). https://doi.org/10.1186/1471-2105-7-97.
16. Hoehndorf, R., Slater, L., Schofield, P.N., Gkoutos, G. V: Aber-OWL: a framework for ontology-based data access in biology. BMC Bioinformatics. 16, 1–9 (2015). https://doi.org/10.1186/s12859-015-0456-9.
17. Vandenbussche, P.-Y., Atemezing, G.A., Poveda-Villalon, M., Vatant, B.: Linked Open Vocabularies (LOV): a gateway to reusable semantic vocabularies on the Web. Semant Web. 1, 1–5 (2014). https://doi.org/10.3233/SW-160213.
18. Till, M., Kutz, O., Codescu, M.: Ontohub: A semantic repository for heterogeneous ontologies. In: Theory Day in Computer Science, DACS'14. p. 2. , Bucharest, Romania (2014).
19. Rueda, C., Bermudez, L., Fredericks, J.: The MMI Ontology Registry and Repository: A Portal for Marine Metadata Interoperability. In: MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges, OCEANS'09. p. 6. , Biloxi, MS, USA (2009).
20. D'Aquin, M., Castro, A.G., Lange, C., Viljanen, K. eds: 1st Workshop on Ontology Repositories and Editors for the Semantic Web, ORES'10. In: 1st Workshop on Ontology Repositories and Editors for the Semantic Web, ORES'10. CEUR-WS.org, Hersonissos, Greece (2010).
21. Jonquet, C., Annane, A., Bouarech, K., Emonet, V., Melzi, S.: SIFR BioPortal : Un portail ouvert et générique d'ontologies et de terminologies biomédicales françaises au service de l'annotation sémantique. In: 16th Journées Francophones d'Informatique Médicale, JFIM'16. p. 16. , Genève, Suisse (2016).
22. Tchechmedjiev, A., Abdaoui, A., Emonet, V., Zevio, S., Jonquet, C.: SIFR Annotator: Ontology-Based Semantic Annotation of French Biomedical Text and Clinical Notes. BMC Bioinformatics. 19, 405–431 (2018). https://doi.org/10.1186/s12859-018-2429-2.
23. Jonquet, C., Toulet, A., Arnaud, E., Aubin, S., Dzalé Yeumo, E., Emonet, V., Graybeal, J., Laporte, M.-A., Musen, M.A., Pesce, V., Larmande, P.: AgroPortal: A vocabulary and ontology repository for agronomy. Comput Electron Agric. 144, 126–143 (2018). https://doi.org/10.1016/j.compag.2017.10.012.
24. D'Aquin, M., Noy, N.F.: Where to Publish and Find Ontologies? A Survey of Ontology Libraries. Web semantics. 11, 96–111 (2012). https://doi.org/10.1016/j.websem.2011.08.005.
25. Naskar, D., Dutta, B.: Ontology Libraries : A Study from an Ontofier and an Ontologist Perspectives. In: 19th International Symposium on Electronic Theses and Dissertations, ETD'16. pp. 1–12. , Lille, France (2016).

Funded by the European Union

26. Ding, L., Finin, T., Joshi, A., Peng, Y., Cost, R.S., Sachs, J., Pan, R., Reddivari, P., Doshi, V.: Swoogle: A Semantic Web Search and Metadata Engine. In: Grossman, D.A., Gravano, L., Zhai, C., Herzog, O., and Evans, D. (eds.) 13th ACM Conference on Information and Knowledge Management, CIKM'04. ACM, Washington DC, USA (2004).
27. D'Aquin, M., Baldassarre, C., Gridinoc, L., Angeletou, S., Sabou, M., Motta, E.: Watson: A Gateway for Next Generation Semantic Web Applications. In: Poster & Demonstration Session at the 6th International Semantic Web Conference, ISWC'07. p. 3. , Busan, Korea (2007).
28. Miles, A., Bechhofer, S.: SKOS simple knowledge organization system reference. W3C recommendation. 18, W3C (2009).
29. Corcho, O., Eriksson, M., Kurowski, K.,, et al.: EOSC interoperability framework. (2021). https://doi.org/https://data.europa.eu/doi/10.2777/620649.
30. Obrst, L., Baclawski, K.: The OOR Initiative – An Update Announcing the new OOR Logo Designed by : Ali Hashemi. 2011, 1–23 (2011).
31. Whetzel, P.L., Team, N.: NCBO Technology: Powering semantically aware applications. Biomedical Semantics. 4S1, 49 (2013). https://doi.org/10.1186/2041-1480-4-S1-S8.
32. Suominen, O., Ylikotila, H., Pessala, S., Lappalainen, M., Frosterus, M., Tuominen, J., Baker, T., Caracciolo, C., Retterath, A.: Publishing SKOS vocabularies with Skosmos. Manuscript submitted for review. (2015).
33. Dai, M., Shah, N.H., Xuan, W., Musen, M.A., Watson, S.J., Athey, B.D., Meng, F.: An Efficient Solution for Mapping Free Text to Ontology Terms. In: American Medical Informatics Association Symposium on Translational BioInformatics, AMIA-TBI'08. , San Francisco, CA, USA (2008).
34. Ghazvinian, A., Noy, N.F., Musen, M.A.: Creating Mappings For Ontologies in Biomedicine: Simple Methods Work. In: American Medical Informatics Association Annual Symposium, AMIA'09. pp. 198–202. , Washington DC, USA (2009).
35. Matentzoglu, N., Balhoff, J.P., Bello, S.M., Bizon, C., Brush, M., Callahan, T.J., Chute, C.G., Duncan, W.D., Evelo, C.T., Gabriel, D., Graybeal, J., Gray, A., Gyori, B.M., Haendel, M., Harmse, H., Harris, N.L., Harrow, I., Hegde, H.B., Hoyt, A.L., Hoyt, C.T., Jiao, D., Jiménez-Ruiz, E., Jupp, S., Kim, H., Koehler, S., Liener, T., Long, Q., Malone, J., McLaughlin, J.A., McMurry, J.A., Moxon, S., Munoz-Torres, M.C., Osumi-Sutherland, D., Overton, J.A., Peters, B., Putman, T., Queralt-Rosinach, N., Shefchek, K., Solbrig, H., Thessen, A., Tudorache, T., Vasilevsky, N., Wagner, A.H., Mungall, C.J.: A Simple Standard for Sharing Ontological Mappings (SSSOM). Database. 2022, (2022). https://doi.org/10.1093/database/baac035.
36. Noy, N.F., Dorf, M., Griffith, N.B., Nyulas, C., Musen, M.A.: Harnessing the Power of the Community in a Library of Biomedical Ontologies. In: Clark, T., Luciano, J.S., Marshall, M.S., Prud'hommeaux, E., and Stephens, S. (eds.) Workshop on Semantic Web Applications in Scientific Discourse, SWASD'09. p. 11. , Washington DC, USA (2009).
37. Jonquet, C., Shah, N.H., Musen, M.A.: The Open Biomedical Annotator. In: American Medical Informatics Association Symposium on Translational BioInformatics, AMIA-TBI'09. pp. 56–60. , San Francisco, CA, USA (2009).
38. Martinez-Romero, M., Jonquet, C., O'Connor, M.J., Graybeal, J., Pazos, A., Musen, M.A.: NCBO Ontology Recommender 2.0: An Enhanced Approach for Biomedical Ontology Recommendation. Biomedical Semantics. 8, (2017). https://doi.org/10.1186/s13326-017-0128-y.
39. Salvadores, M., Horridge, M., Alexander, P.R., Fergerson, R.W., Musen, M.A., Noy, N.F.N.F.: Using SPARQL to query bioportal ontologies and metadata. In: Cudr&#65533;-Mauroux, P., Heflin, J., Sirin, E., Tudorache, T., Euzenat, J., Hauswirth, M., Parreira, J.X., Hendler, J., Schreiber, G., and Bernstein, A. (eds.) 11th International Semantic Web Conference, ISWC'12. pp. 180–195. Springer, Boston, MA, USA (2012). https://doi.org/10.1007/978-3-642-35173-0_12.
40. Melzi, S., Jonquet, C.: Scoring semantic annotations returned by the NCBO Annotator. In: Paschke, A., Burger, A., Romano, P., Marshall, M.S., and Splendiani, A. (eds.) 7th International Semantic Web Applications and Tools for Life Sciences, SWAT4LS'14. p. 15. CEUR-WS.org, Berlin, Germany (2014).
41. Dutta, B., Toulet, A., Emonet, V., Jonquet, C.: New Generation Metadata vocabulary for Ontology Description and Publication. In: Garoufallou, E., Virkus, S., and Alemu, G. (eds.) 11th Metadata and Semantics Research Conference, MTSR'17. , Tallinn, Estonia (2017). https://doi.org/10.1007/978-3-319-70863-8_17.
42. Amdouni, E., Bouazzouni, S., Jonquet, C.: O'FAIRe: Ontology FAIRness Evaluator in the AgroPortal Semantic Resource Repository. In: 19th Extended Semantic Web Conference, ESWC'22, Demo and

Funded by the European Union

Poster session. pp. 89–94. LNCS, Springer, Hersonissos, Crete, Greece (2022). https://doi.org/10.1007/978-3-031-11609-4_17.
43. Jonquet, C., Bouazzouni, S.: State-of-the-art browsing of SKOS resources in AgroPortal. In: NKOS Consolidated Workshop. , Daegu, South Korea (2023).
44. Jonquet, C., Graybeal, J., OntoPortal Alliance, M. of the: OntoPortal Workshop 2022 Report. , Montpellier, France (2022).
45. Jonquet, C.: Ontology Repository and Ontology-Based Services – Challenges, contributions and applications to biomedicine & agronomy, https://tel.archives-ouvertes.fr/, (2019).

Funded by the European Union