Serving FAIR Semantic Artefacts: A Review of Semantic Artefact Catalogues

Clement Jonquet,[1] Nina Grau,[1] Guillaume Alviset,[2] Baptiste Cecconi, [3] Daniel Garijo,[4] María Poveda-Villalón,[4] Martina Pulieri, [5] Ilaria Rosati[6],[7]

Abstract

The growing number of ontologies and other semantic artefacts in data-intensive research has made their discovery, evaluation, and reuse a critical challenge. Semantic Artefact Catalogues (SACs) have emerged as key infrastructures to address these needs, yet their landscape seems fragmented and poorly characterised. Questions persist regarding which types of SACs exist (ontology repositories/libraries, vocabulary/terminology services, etc.), which are actively maintained, the domains they serve, and the technologies on which they are based. This paper presents a comprehensive current and historical review of SACs. The study focuses on their technologies, disciplinary scope, and capacity to enable FAIR principles for semantic artefacts. 190 SACs were identified and analysed, across 17 disciplines and relying on 12 “generic” technologies. To assess their FAIR-enabling potential, we defined ten dimensions derived from existing ontology FAIRness assessment frameworks and tools and assessed a subset of 60 SACs accordingly. The findings underscore both the diversity and fragmentation of SACs, while identifying opportunities for greater harmonisation and interoperability through shared, reusable technologies or APIs. A living version of the SAC review dataset is openly available (DOI: 10.5281/zenodo.12799861) and can be continuously updated or extended by the community to reflect new catalogues and developments.

Keywords

Semantic Artefacts, Ontologies, Vocabularies, Terminologies, Semantic Artefact Catalogues, Ontology repositories, Vocabulary services, Terminology servers

Introduction

Semantic Artefact (SA) is a broad term used to designate a wide range of knowledge organisation systems including ontologies, terminologies, taxonomies, thesauri, vocabularies, and metadata schemas [1][2]. Acting as shared conceptualisations of knowledge, these artefacts are essential to standardising data representation and annotation. Across all scientific disciplines, SAs are extensively utilised to represent and annotate data in a standardised way. They encapsulate high-level domain knowledge and play a central role in ensuring interoperability across datasets,

[1] MISTEA, Univ. of Montpellier, INRAE & Institut Agro, Montpellier, France.
[2] CNRS Data Terra, France.
[3] LIRA, Observatoire de Paris, Univ. PSL, CNRS, Sorbonne Univ., Univ. Paris Cité, Meudon, France.
[4] Ontology Engineering Group, Universidad Politécnica de Madrid, Madrid, Spain
[5] University of Palermo, Italy.
[6] IRET, CNR, Lecce, Italy.
[7] LifeWatch Italy, Lecce, Italy.

systems, and disciplines such as for instance in environmental sciences [3].[8] Ultimately SAs support the FAIR Data Principles [4], particularly Principle I.2 which focus on enabling Interoperability through the use of FAIR vocabularies, and they are increasingly recognised as research objects that must also themselves comply with FAIR principles.

In the rapidly evolving landscape of scientific research, the proliferation of semantic artefacts requires the development of robust systems to manage and utilise these resources effectively [5],[6]. To maximise the discoverability, accessibility, and reuse of SAs, it is crucial to organise and share them through Semantic Artefact Catalogues (SACs), structured and accessible platforms that include ontology repositories, libraries, listings, registries, hubs or portals. SACs provide essential platforms for receiving, hosting, serving, aligning, and enabling the reuse of SAs, playing a pivotal role in the identification, description, evaluation, and dissemination of SAs. SACs are therefore key enablers for making semantic artefacts themselves FAIR digital objects, becoming essential components of the SA lifecycle.

SAs and SACs are becoming first class citizens for Research Infrastructures as well. Both are acknowledged within the European Open Science Cloud (EOSC) Interoperability Framework [7] as core technical components and studied by the task forces [2] and projects [1][8]. A growing number of SACs have emerged across different domains and communities, often developed independently with domain-specific priorities, different technology stacks and design principles, leading to fragmentation and interoperability challenges. In addition, approaches to SAC interoperability are now being studied and we start seeing results to federate [9] or interconnect them with shared APIs [10].

The types of semantic artefact catalogues ranges from simple SA listings to rich libraries (such as the OBO Foundry, the BARTOC registry) with structured metadata, and to advanced repositories (or portals) that offer a variety of services for multiple types of semantic artefacts (such as OntoPortal instances, OLS, or the LOV platform). These services may include browsing/searching, visualisation, metrics, recommendations, and annotation of data. SACs are often developed or maintained by specific discipline communities or infrastructures and we have seen the emergence of specific generic technologies that can be used to deploy new semantic artefact catalogues (such as OntoPortal, OLS or Skosmos). D'Aquin & Noy [11] provided in 2012 a reference review of ontology repositories at that time. Later completed by Naskar and Dutta in 2016 [12]. However, these two studies, now largely outdated and focused on ontologies, required another comprehensive and rigorous update.

In this article, we explore the current landscape of SACs; we make a comprehensive review of current (and past) SACs that can be used to search and host SAs. This study was originally made within the European project FAIR-IMPACT [13], one project of the EOSC ecosystem, to assess SACs current state and historical development, including the technologies used, the type of SAC found, and the communities and themes covered. We successfully collected information on 190 SACs, of which 129 are active (including 9 prototypes), across 17 disciplines and 12 "generic" technologies. Plus, based on five methodologies and tools for FAIRness assessment of SAs, (O'FAIRe, FOOPS!, FsF, 10-SR and FVF – detailed and referenced later) we have regrouped 10

[8] According to the FAIRsFAIR project [1] and the EOSC Interoperability Framework (EOSC-IF) [7] which first introduced the expression, SAs are defined as "machine-actionable and readable formalization of a conceptualisation that enables the sharing and reuse by humans and machines. These artefacts may have a broad range of formalisation, from loose set of terms, taxonomies, thesauri to higher-order logics."

key dimensions for assessing FAIR-enabling capabilities of SACs. For almost half of the active SAC (60) we studied how much each SAC enables or supports FAIR for their artefacts. Therefore, this review provides a good overview of how SACs can help/support SAs to address FAIR principles and contribute to the efficient management and utilisation of SAs.

In the rest of the paper, we first introduce the definitions and context of this research, followed by an historical perspective on SACs in Section 2 which redraws the related academic work. Section 3 presents the methodology used for a comprehensive review of SAC and describes how we classified SAC based on their status, types, and technology used. This section also presents the FAIR-enabling dimensions used to evaluate how integrating SA into SAC enhance their FAIRness. Section 4 presents our results and analysis, with a state-of-the-art list of SACs gathered, reviewed and assessed against the FAIR-enabling dimensions. We conclude in Section 5.

The associated data of this review are publicly available under the form of a spreadsheet which contains the listing of SACs, their classifications (status, type, discipline, technology) and the FAIR-enabling assessment (https://doi.org/10.5281/zenodo.12799861). The spreadsheet is versioned and open for changes to the community to reflect new catalogues and developments.

Background

From ontology libraries and repositories to semantic artefact catalogues[9]

The continuous growth in the number of ontologies has made ontology libraries and repositories a longstanding topic of interest in the Semantic Web community. Ding & Fensel [14] introduced the notion of *ontology library* in 2001 and presented a review of libraries at that time:

> "A library system that offers various functions for managing, adapting and standardizing groups of ontologies. It should fulfill the needs for re-use of ontologies. In this sense, an ontology library system should be easily accessible and offer efficient support for re-using existing relevant ontologies and standardizing them based on upper-level ontologies and ontology representation languages."

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[10]) or by offering structured metadata to describe them (e.g., BARTOC [15], Agrisemantics Map [16]). But those systems generally do not support any services beyond description, including services based on the content of the ontologies. In the biomedical domain, the OBO Foundry [17] is a reference library effort to help the biomedical and biological communities build their ontologies with an enforcement of design and reuse principles.

[9] This section on historical perspective on ontology repositories and SACs relies and completes a section in [30], which introduced related works before presenting the OntoPortal technology. This section was itself inspired from C. Jonquet's habilitation thesis [6].

[10] Cf: https://www.daml.org/ontologies/ other listings are now unavailable anymore.

FAIRsharing [18] registers many SAs among multiple other data standards. Other historical libraries include [19][20].

Hartman et al. [21] introduced in 2009 the concept of *ontology repository* with advanced features based on the content of the ontologies such as search, browsing, metadata management, visualization, personalization, mappings and an application programming interface to query the repository's content and services:

> "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. Therefore, a repository provides a formal query language."

Among the first and more prominent ontology services we can cite the NCBO BioPortal [22] but other early initiatives discontinued today include OntoServer [23], OntoSelect [24], OntoSearch [25, 26] or CupBoard [27]. By the end of the 2000's, the topic was of high interest as illustrated by the 2010 ORES workshop [28] or the 2008 Ontology Summit 2008.[11] The Open Ontology Repository (OOR) Initiative [29] was a collaborative effort to develop a federated infrastructure of ontology repositories. At that time, the effort already reused the BioPortal technology that was the most advanced open-source technology for managing ontologies but not yet packaged in a "virtual appliance" as it is today with OntoPortal [30]. Later the OOR initiative studied OntoHub [31] technology for generalization but this initiative is now discontinued.

In parallel, there have been efforts to index any semantic web data online (including ontologies) and offer search engines such as Swoogle [32] and Watson [33]. We cannot talk about ontology libraries or repositories for those "semantic web indexes", even if they support some features of ontology repositories (e.g., search).

Other similar products that match as repositories were called terminology services or vocabulary servers and were usually developed to host one or a few terminologies for a specific community (e.g., SNOMED-CT terminology server, UMLS-KS, CLARIN vocabulary services); they are usually not semantic web compliant and do not handle the complexity of ontologies. The SKOS specification in 2009, inspired a spike in the number of development of tools compliant with or tailored for this representation language, of which Skosmos [34] is the best example as it was developed as a technology to be easily reused for dedicated vocabulary servers (although it only supports browsing and search for SKOS vocabularies). For instances, Finto [35] or Loterre[12] have adopted and customised Skosmos as backend technology.

Multiple ontologies repositories have been developed since the end of the 2010s. The biomedical domain has seen the development of lot of applications (not necessarily synchronised) besides the NCBO BioPortal [22], were built OntoBee [36], the EBI Ontology Lookup Service [37] and AberOWL [38] or the first multilingual repository, HeTOP, the Health Terminology Ontology Portal [39]. We have seen also repository initiatives such as the Linked Open Vocabularies [40], OntoHub [31], and the Marine Metadata Initiative's Ontology Registry and Repository [41] and its earth science counterpart, the ESIP Federation's Community Ontology Repository. The two

[11] https://ontologforum.org/index.php/OntologySummit2008

[12] https://loterre.istex.fr

reviews previously mentioned may include additional historical work: D'Aquin & Noy [11] in 2012 and Naskar and Dutta in 2016 [12].

Many approaches have been proposed throughout the years, leading to fragmentation, duplication and making them difficult to sustain. This "hype" period (early 2010s) was characterised by the creation of a brand-new technology each time a new repository was built which stimulated innovation and new ideas (e.g., where BioPortal was based on an RDF triple-store backend, OLS implemented property graph storage; AberOWL took OWL reasoning at its core, whereas this aspect was ignored by the other approaches) but ended up duplicating many development efforts and has created some confusion in some communities. Typically, the biomedical domain, which saw the most innovative approaches also struggled with the more confusion leaving final users to work with multiple heterogeneous repositories not synchronised, nor managed in a collaborative way. This luxury of developing too many solutions happened though mostly in the biomedical domain. Multiple relevant abandoned projects/tools relevant to the subject include: Cupboard, Knoodl, Schemapedia, SchemaWeb, OntoSelect, OntoSearch, OntoSearch2, TONES, SchemaCache, Soboleo, LexEVS (Mayo Clinic), Intelligent Topic Manager (Mondeca).

The expression Semantic Artefact Catalogues (SACs) was introduced by the European Open Science Cloud (EOSC) Interoperability Framework [7] and then later adopted in reviews [2] or technology paper [30] or multiple EU projects [1, 8]. This expression translates the idea that such catalogues are not only for ontologies or semantically rich vocabularies but must offer common services for a wide range of semantic artefacts. We here define it as an overarching term for the other expressions introduced before.

Semantic Artefacts Catalogues

The inclusion (or deposit) of data in an "open repository came out as the most important factor when determining the quality of a dataset" [42] and the role of data repositories to support FAIR (clearly established by Principle F4) [4] is now considered obvious. In this paper, the data repositories of interest are Semantic Artefact Catalogues (SACs).

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. Besides offering SA developers a platform to share their artefacts, SACs assist general users in handling and using SAs without requiring them to manage the complex and time-consuming process of developing SA themselves. Additionally, like any other data repositories, they play a crucial role in making the SAs they host compliant with the FAIR principles. Our goal is to establish a comparison framework to ease the selection and use of SACs.

Semantic Artefacts Catalogue technologies

In terms of reusable technologies, the NCBO BioPortal [22] 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 OOR Initiative, National Cancer Institute, and Marine Metadata Initiative), 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, i.e., 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 from 2018 [30]. In 2015, the SIFR BioPortal [43] prototype was built to develop a French Annotator and experiment with multilingual issues in BioPortal [44]. 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 [45]. This new approach, along with the Skosmos philosophy [34], to build repositories based on an already existing technology was sometimes later adopted in 2022 by TIB that have reused the OLS technology [37].[13] Another example is VocPrez, an open-source technology developed by a company adopted for example by the Geoscience Australia Vocabularies system[14] or by the NERC Vocabulary Server [46]. Also, we can cite ShowVoc [47], a recently developed technology for SACs based on the same core components as VocBench vocabulary editor [48] but it appears that Shovoc have drawn most of its inspiration from OntoPortal in terms of its design and services. Since 2020, the OntoPortal technology has been reused in multiple other communities from ecology (with EcoPortal) [49], earth sciences (with EarthPortal) [50], biodiversity (with BiodivPortal) [51], astronomy (with OntoPortal-Astro) [52] and more.

EOSC: FAIR data requires FAIR semantic artefacts and their catalogues

FAIR-IMPACT was a project in the EOSC program with a dedicated focus on semantic artefacts and their catalogues. It worked with multiple scientific communities to consolidate, deploy or experiment SACs for their disciplines (AgroPortal EcoPortal, EarthPortal, OntoPortal-Astro). FAIR-IMPACT has also consolidated the MOD (Metadata for Ontology Description and Publication Ontology), a DCAT2-based standard vocabulary to describe semantic artefacts (https://github.com/FAIR-IMPACT/MOD) and have published a shared API for semantic artefact catalogues based on MOD (https://github.com/FAIR-IMPACT/MOD-API) that has/is been implemented by multiple SACs originally gathered in a FAIR-IMPACT funded action.

In parallel, the EOSC Semantic Interoperability Task Force (TF) (finished end of 2023) also had a dedicated topic on semantic artefact ("*a critical part of the long-term viability of any research data infrastructure*") and did produce results related to SACs [53]: 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 [2]. This model includes 12 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. Our work is complementary to the EOSC TF effort in two main aspects: (i) our review of SACs is larger (7 times more SAC reviewed) 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.

Methodology

The growing centrality of ontologies and other semantic artefacts in data-intensive research raises fundamental questions about their discovery, evaluation, and reuse. Once an ontology is created, how can it be effectively disseminated and made visible to relevant communities? Conversely, when an ontology is needed, where should it be found, and how can its quality, relevance, and

[13] https://terminology.tib.eu
[14] https://vocabs.ardc.edu.au

fitness for purpose be assessed? Semantic Artefact Catalogues (SACs) have emerged as key infrastructures to address these questions. But what types of SACs are available now for use? Which ones are still active? In which domains? Are they based on a generic or ad-hoc technology? This section presents the methodology adopted to systematically review and analyse SACs, with the objective of understanding how they support these needs and enable FAIR, reusable semantic artefacts across disciplines. In this section we first present our methodology on building a comprehensive listing of semantic artefact catalogues and how we classified them. We then explain the FAIR-enabling dimensions and assessment of the SACs.

Comprehensive listing of Semantic Artefact Catalogues

Related work and listing process

For our review, SACs were collected from diverse sources, categorised by type, discipline and technology. The authors –each of whom is a provider of at least one Semantic Artefact Catalogue for their respective community– together with a small number of additional FAIR-IMPACT members (including regular semantic artefact users, developers, and data stakeholders with expertise in semantics), totalling approximately a dozen contributors, were initially asked to identify and list any SACs or collections of SACs they were aware of in a shared document. The list was also populated with resources –collected by desk research and experience/expertise of the members– and identified within:

- C. Jonquet's Habilitation Background chapter [6];
- Noy and d'Aquin 2012's historical article [11];
- Naskar and Dutta's article based on an internal ISI report in 2016 [12];
- Work from the EOSC Semantic Interoperability TF theme 2 on establishing a maturity model for SAC [2];
- The state-of-the-art and Section 2 of the article presenting the OntoPortal technology [30];
- The BARTOC Terminology Registries list which did a similar and quite rich listing.[15]

Then we also analysed the homepages and references of the SAC technologies already known or identified during our review especially: OntoPortal, Skosmos, TemaTres [54] and OpenTheso[16] which usefully listed the different installations of their technology. In addition, we have explicitly contacted the providers of each of the technologies, asking them for an exhaustive listing of the installations of their technologies, if such a listing was available and public.

Once all individual SAC were listed, the two first authors transferred them into a shared working spreadsheet qualifying them, removing the duplicates, and compiling their descriptive information as described below. Once all fields of the table were completed, we performed a double-checking strategy to verify the information with one of the two first authors controlling the information of the other and doing another loop with the rest of the authors controlling some of the content (usually the SACs they were mostly familiar with or responsible for).

[15] https://bartoc.org/registries Approximately 30% of the SAC reviewed were sourced from this list (i.e., not found on any other previously listed sources), which significantly enriched our SAC collection.
[16] https://opentheso.huma-num.fr

General information

For each SAC reviewed, we identified the **name** and current web application **URL** 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).

Status

The goal of our review was to establish a state-of-the-art listing of what catalogues are available now for use. However, we decided to include relevant catalogues that may have been discontinued as well. In this sense, our review also serves as a historical record of the SAC landscape. This aspect is represented by the **status**, mostly determined based on the maintenance/availability of its URL or other recent information or changes to the application. A SAC is considered "Retired" if the URL is no longer accessible and/or functional at the time of review, 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.

Types of Semantic Artefact Catalogues

Based on the definitions above, we view a SAC as a web platform where multiple (more than one) SAs are at minimum listed or indexed, but in most cases stored, preserved, served and shared. We have grouped catalogues in the following **types**:

- *Listing*: A simple online listing of semantic artefacts on a web page or 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. Listings are typically implemented by means of a simple HTML page.
- *Library*: 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. Libraries 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. Libraries are almost always based on some ad-hoc technology.
- *Search Service*: Online index that is specifically dedicated to indexing/searching semantic artefacts. These services do not host or serve directly the content of artefacts, but as any other search engine, it offers means to identify SA content and link to it.
- *Repository*: Advanced web application with 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/or automatic generation, text annotation or recommendation of SAs. Repositories can be queried by machine via APIs or SPARQL endpoints. Repositories are often based on a common, generic technology.
- *Other*: SACs that do not fit with the previously defined types but that deserve to be listed. In our report, we identified only three SACs as "Other."

Disciplines

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

- In a folksonomy approach, the first step was conducted by the authors who categorised each SAC with "open" keywords corresponding to disciplines, based on their knowledge of the listed SACs, their content, and the descriptions provided. This process yielded 37 unharmonized disciplines, with a 10% representation of unique disciplines. And 89 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 disciplines and ease comparison and analysis, we decided to align them with an already established domain classification. We used the list of academic disciplines provided by Wikipedia to which it's easy to refer and well described.[17] To achieve this alignment, we used a large language model (i.e., ChatGPT[18]) to pre-classify the SAC from our collection based on their Web URL. The prompt used was: "*According to the list proposed by Wikipedia: (above footnote), classify by field of 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. We refined the classification to 17 disciplines standardised in Wikipedia, with less than 5% representation of unique disciplines. We use only these 17 disciplines plus the "General" category in our analysis.

Technology used

We identified whether a target SAC was running on a **generic technology** that is (or may be) used to deploy other catalogues (e.g., in different projects or disciplines). The SAC technology was determined by our knowledge of this area and acquaintance with common technologies available to deploy or install a new SAC. Typically, a technology used to set up a SAC was considered "generic" if we were able to find a description of this technology, a website and/or source code repository, or whether we were able to find multiple SACs running on the same technology.

The different SAC technology identified have been listed in a separate tab of our spreadsheet; for each we identified (cf. Table 1): 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).

Table 1. List of Semantic Artefact Catalogue technologies.

[17] Wikipedia of academic fields. URL: https://en.wikipedia.org/wiki/List_of_academic_fields
[18] OpenAI, GPT-4 model 2024 (free version). URL: https://chatgpt.com

Name	Related URL	Open source	Code repository	Status	Contact
OntoPortal	https://ontoportal.org/	Yes	https://github.com/ontoportal	Active	Clement Jonquet
OLS	https://github.com/EBISPOT/ols4	Yes	https://github.com/EBISPOT/ols4	Active	Henriette Harmse
Skosmos	https://skosmos.org/	Yes	https://github.com/NatLibFi/Skosmos	Active	Osma Suominen
ShowVoc	https://showvoc.uniroma2.it/	Yes	https://bitbucket.org/art-uniroma2/showvoc/src/master/	Active	Armando Stellato
TemaTres	https://vocabularyserver.com/web/about	Yes	https://github.com/tematres/TemaTres-Vocabulary-Server/	Active	Diego Ferreyra
ONKI SKOS Server	https://seco.cs.aalto.fi/services/onkiskos/	Yes	https://seco.cs.aalto.fi/services/onkiskos/onki-skos-20121221.zip	Retired	Jouni Tuominen
Vocab	https://github.com/oeg-upm/vocab.linkeddata.es	Yes	https://github.com/oeg-upm/vocab.linkeddata.es	Active	María Poveda Daniel Garijo
OpenTheso	https://opentheso.hypotheses.org/	Yes	https://github.com/miledrousset/Opentheso2	Active	Miled Rousset
iQvoc	https://iqvoc.net/	Yes	https://github.com/innoq/iqvoc	Active	info@innoq.com
ORR	https://mmisw.org/orrdoc/about/	Yes	https://github.com/mmisw	Active	Carlos Rueda
VocPrez	https://rdflib.dev/VocPrez/	Yes	https://github.com/RDFLib/VocPrez/	Retired	Nicholas Car
Prez	https://github.com/RDFLib/Prez	Yes	https://github.com/RDFLib/Prez	Active	Nicholas Car
Centree	https://scibite.com/platform/centree-ontology-management-platform/	No	NA	Active	Simon Jupp
Termweb	https://www.interverbumtech.com/products-services/termweb/	No	NA	Active	info.eu@interverbumtech.com
Data Harmony Taxonomy Suite	https://www.accessinn.com/data-harmony-products/	No	NA	Active	https://www.accessinn.com/contact-us/
OCLC Terminology Services	http://tspilot.oclc.org/resources/	No	NA	Retired	oclc@oclc.org
OpenSKOS	http://openskos.org/	No	NA	Retired	CATCHPlus
THEMAS	https://www.ics.forth.gr/isl/themas-thesaurus-management-system	Yes	https://github.com/isl/THEMAS	Retired	Elias Tzortzakakis
SkoHub	https://skohub.io/	Yes	https://github.com/skohub-io	Active	skohub@hbz-nrw.de
SKOS Shuttle	https://semweb.solutions/skosshuttle/	No	NA	Active	Fabio Ricci

Exclusion criteria

We decide not to include in the list of SACs tools, websites or applications (vocabulary services or terminology service) with only one semantic artefact. Many ad-hoc web applications were developed to host and serve one unique SA 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 SACs exposing only a single semantic artefact were not exhaustively reviewed or analysed in this study, some instances encountered during the systematic review were nonetheless retained and recorded separately in a dedicated tab of the accompanying spreadsheet. Those are not counted nor presented in the results. 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 in our collected data (but not the list of SACs). Some examples (Bioschemas, LusTRE, OKFN, etc.) are available in dedicated tabs of the spreadsheet associated with this study.

FAIR-enabling dimensions and SAC FAIR-enabling assessments

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. Here, 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. In order to produce a set of FAIR-enabling dimensions for SACs, we reviewed the five state-of-the-art methods or tools to evaluate/assess the level of FAIRness of semantic artefacts. These tools and methods are (order more or less chronologically):

- **FAIRsFAIR's Recommendations (FsF P-Recs) [1]** were produced within the eponym H2020 project. It's a list of 17 recommendations and 10 best practices 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.
- **FOOPS! (Ontology Pitfall Scanner for the FAIR principles) [56]** 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 detects best practices according to each principle, but also offers an explanation of why a particular principle fails, and helpful suggestions to overcome common issues.
- **O'FAIRe (Ontology FAIRness Evaluator) [55]** a methodology and tool designed to assess the FAIRness of ontologies and semantic artefacts, based on 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 OntoPortal-based SACs, 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.
- **10-simple rules (Ten simple rules for making a vocabulary FAIR) [57]** 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) [58]** 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 and 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, criteria of these tools and methods depend 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. Similarly, some recommendations of FAIRsFAIR are directly addressed to the catalogues (P-Rec 5 or 7).

We reviewed the criteria, questions or recommendations of the five tools and methods listed above and regrouped them in **10 broad dimensions.** This work was not an alignment or uniformization of the five tools and methods which each keep their diversity and specificity in assessing FAIRness, but more of a broader prism to uniformly browse them. The 10 broad dimensions are:

- *Identifiers*. Includes criteria related to SA's persistent and resolvable identifiers. 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*. Includes criteria related to SA metadata e.g., follows the MIRO guidelines [59], including mandatory properties like title, description, licence, and creation date. Additional metadata properties are encouraged too. Metadata may 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, properties, definitions, hierarchy, labels, etc. Provenance information and a consistent schema are essential for maintaining the SA's integrity and usability.
- *Cataloguing*. Includes criteria related to public registration of the SA in multiple SACs, ensuring wide accessibility and indexing by web search engines. Public registration in trustworthy catalogues enhances discoverability and maintenance traceability. Indexed SA are easier to search, find, and use, promoting broader adoption and interoperability.
- *Resolvability*. Includes criteria 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*. Includes criteria 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*. Includes criteria 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*. Includes criteria related to the encoding of the SA that must use standardised representation languages, preferably those recommended by the W3C (e.g., RDFS, OWL, SKOS). The representation language, syntax, and formality level should be clearly informed. Availability in multiple syntaxes (TTL, RDF/XML, etc.) 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*. Includes criteria 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*. Includes criteria 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*. Includes criteria related to SA provenance information, like actors involved, accrual methods, versioning, latest changes, methodology, tools, and 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 criteria, questions and recommendations available in each dimension.

FAIR-enabling levels

For each SAC identified, we determined how much a SAC supports or enables the criteria 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 2.

Table 2. SAC FAIR-enabling level with respect to the FAIR-enabling dimensions.

Level 3	The SAC mostly handles this dimension (either automatically or through manual check). SA developers do not have anything else to do.
Level 2	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 criteria.
Level 1	The SAC mostly does not support this dimension at all. It's totally independent of the SAC for SA developers to address these criteria.
Level 0	The role of the SAC with respect to this dimension has not been assessed (is hard to determine (not apply to the catalogues) or its unknown).

FAIR-enabling analysis methodology

Our FAIR-enabling assessment was intentionally kept at a coarse-grained level, focusing on high-level capabilities rather than detailed compliance. Our aim was not to compare the SACs at a fine-grained level, but to show which dimensions are broadly addressed by which type of SACs, where a SAC may have to put efforts in the future towards addressing gaps, and to emphasise specific aspects of certain SACs so they could serve as an example to other SACs. We applied and mixed multiple strategies to assess the FAIR-enabling level of the SACs in our study:

- The following repositories were self-assessed by their maintainers or developers (within the authors) and the assessment was discussed and harmonised during meetings: SIFR BioPortal (INRAE), AgroPortal (INRAE), EcoPortal (LifeWatch), EarthPortal (Data

Terra), LOV (UPM), Identifiers.org (UNIMAN), OEG Vocabularies (UPM), IVOA Vocabularies (IVOA).

- 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 "by default" based on our task members knowledge of the technology. This applied to SACs based on the following technologies: OLS, ORR, ShowVoc, VocPrez, and to some extent OntoPortal instances not assessed by their maintainers (first bullet point).
- Each time we were aware of a particular enforcement or attention a SAC would pay a specific attention to; we changed the default 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 "Level 2" (orange) (as other OntoPortals) to to "Level 3" (green) in EcoPortal's case.
- Prototypes and retired SACs were removed from the study.

The FAIR-enabling assessment was conducted on a subset of 60 out of the 129 active SACs (46%). This subset was limited to catalogues for which the authors had direct expertise, either as providers, maintainers, or long-term users. The objective of this assessment was not to be exhaustive nor to rank or benchmark existing SACs, but rather to derive representative insights at a coarse-grained level regarding how different types of catalogues and technologies support FAIR-enabling dimensions. Given the interpretative and partially subjective nature of such an assessment at this level of abstraction, we restricted the analysis to catalogues we knew well, in order to minimise misinterpretation. In line with the overall philosophy of this study, the underlying dataset and FAIR-enabling assessments are maintained as an open, living document that can be updated, refined, or extended over time, including through contributions from SAC providers and users beyond the author group.

Results

The complete list of all the SAC studied was published in the version 1 of the associated spreadsheet data (DOI:10.5281/zenodo.12799862) in July 2024. With feedback and information contributed by external parties since then, with this article, we produced an updated version of the list (DOI: 10.5281/zenodo.12799861). A total of 190 semantic artefact catalogues were included in our review.

The following section (except next one) 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 customisable, informative reports and dashboards from *Google Sheets*.

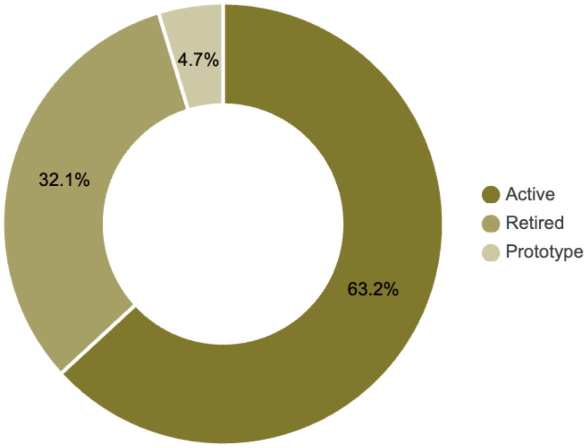
Comprehensive listing of Semantic Artefact Catalogues

Status

Over 120 "Active", 61 "Retired", and 9 "Prototype" catalogues were reviewed (Fig. 1). This indicates that a significant proportion (63%) of the gathered SACs are currently usable and the important number of SACs (120) demonstrates how relevant and necessary they are. Among the 32% of the “Retired” SACs a few 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 ▾
Active	120
Retired	61
Prototype	9

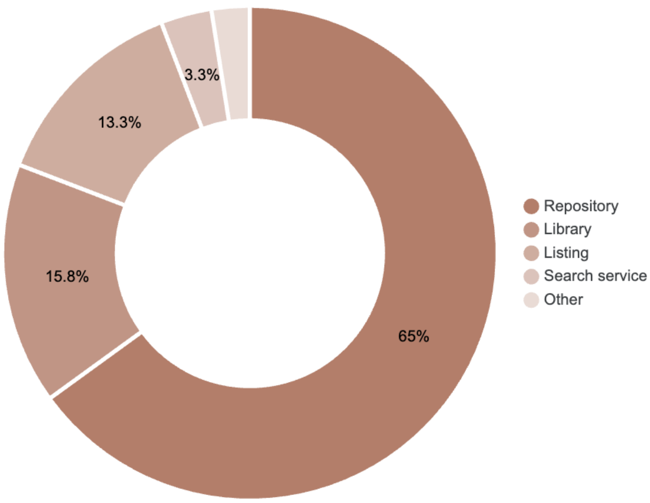


Types of SACs

Among the 120 “Active” SAC” collected, the “Repository” type is the most prevalent, representing nearly 78% of the reviewed SACs (Fig. 2). The “Library” and “Listing” types are similarly represented in our listing, comprising 19% and 16% of the SAC, respectively. Finally, the “Search Service” type is the least represented, with only 4 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 (78) 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.

Type	Record Count ▾
Repository	78
Library	19
Listing	16
Search service	4
Other	3



Generic technology

Regarding the technology used by the active SACs, we identified 20 different technologies (Table 1) but found running instances of only 12 of these technologies (Fig. 3), because:

- 3 private/non-open technologies out of 5 (Centree, OCLC Terminology Services) did not make public the use of their technologies (usually their business clients) or were obsolete (OpenSKOS, VocPrez).[19]
- Prez and VocPrez were joined and counted 1 (VocPrez being retired and replaced by Prez).
- 3 technologies were identified but without catalogues: THEMAS, ONKI SKOS Server (retired), SkoHub (unclear).

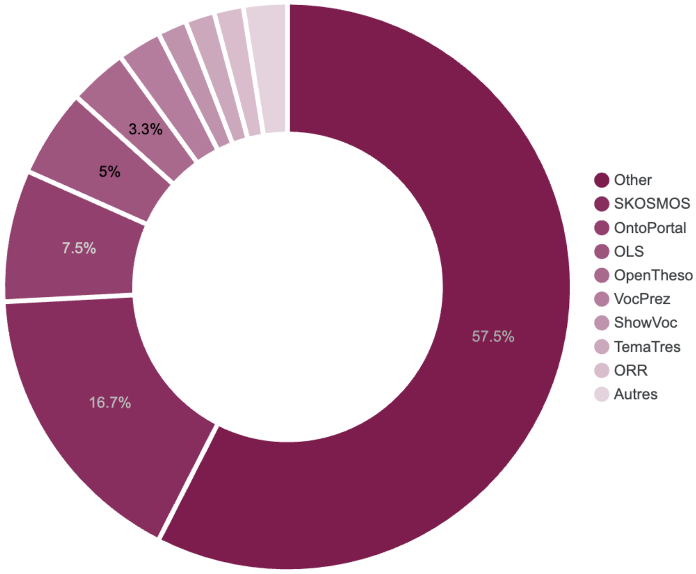
When we were unable to determine the technology (and its genericity), we categorised the undefined ones in the "Other" category which represents more than half (57%) of the SACs.

The top three most prevalent known technologies in our list are Skosmos, OntoPortal and OLS, which represent respectively 16.7%, 7.5% and 5% 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 most focused technology and easy to install, which indirectly explains the fact that it's the most reused [34] (https://skosmos.org). 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 Skosmos based SACs never accept direct SA submission from external parties. We may also observe that the OntoPortal technology [30] (https://ontoportal.org), 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 four of the reuses (over five) of the OLS technology are done by one big infrastructure program (NFDI via TIB) showing less diversity of reuses than OntoPortal (https://terminology.tib.eu). Other observations related to the use of the OntoPortal technology are reported in [9, 30] where the authors argue one of the main reasons for adopting this technology is to join an active community and Alliance providing explicitly engaged in enabling and supporting the re-use of its technology; a commitment we have not observed in any other open-source technology.

Figure 3. Number and distribution of all SAC listed with active status by generic technology used. Note that if prototype SAC would have been included OntoPortal would record 4 more, Skosmos 2 more and ShowVoc 1 more.

[19] Note that we have identified 5 private technologies but more may exist. This is beyond the scope of this work to review private or business-driven solutions.

	Generic technology	Record Count ▾
1.	Other	69
2.	SKOSMOS	20
3.	OntoPortal	9
4.	OLS	6
5.	OpenTheso	4
6.	VocPrez	3
7.	ShowVoc	2
8.	TemaTres	2
9.	ORR	2
10.	iQvoc	1
11.	Vocab	1
12.	Termweb	1



Disciplines

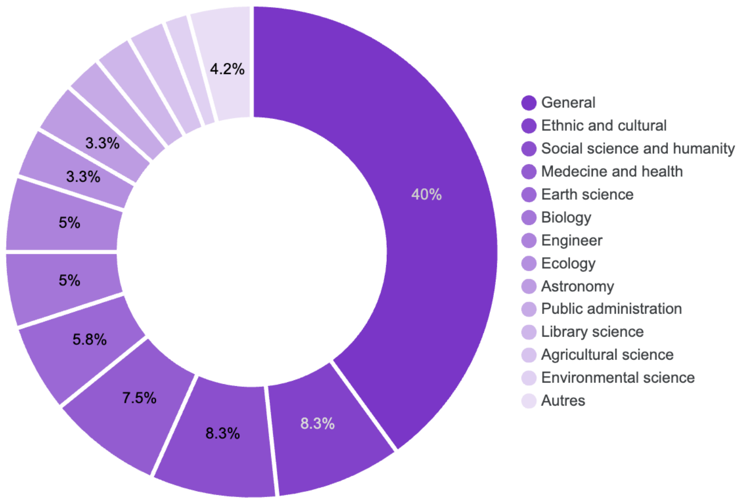
We categorised 17 disciplines plus the “General” category in our analysis (for further more details, refer to the Discipline sub-section of the Methodology section). The “General” category accounts for 40% of active SACs (Fig. 4) that are not domain specific thus are targeting users from various communities 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 active SACs. A second gap is observed among disciplines that have 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”, “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 in this study. If “Ecology” were included within the “Biology” field, “Biology” would become the second most represented domain with 10 SAC. Another observation is the prominence of the “Medicine and health” and “Biology” disciplines that together make 14 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 standardized Wikipedia disciplines.

	Standardized Wikipedia Disci...	Record Count ▾
1.	General	48
2.	Ethnic and cultural	10
3.	Social science and humanity	10
4.	Medecine and health	9
5.	Earth science	7
6.	Biology	6
7.	Engineer	6
8.	Ecology	4
9.	Astronomy	4
10.	Library science	3
11.	Public administration	3
12.	Agricultural science	3
13.	Environmental science	2
14.	Linguistic	1
15.	Economics	1
16.	Material sciences	1
17.	Education	1
18.	Chemistry	1



Cross-analysis of disciplines 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 is 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 *Listing*. "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 SACs: *Listing* and *Library*, without including the *Repository* type. Finally, there is only one discipline - “Astronomy” - that uses three types of SACs, which are *Repository*, *Listing* and *Library*, for historical reasons.

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

	Type / Record Count					
Standardized Wikipedia Discipline	Repository	Library	Listing	Search service	Other	Total...
Agricultural science	3	-	-	-	-	3
Astronomy	2	1	1	-	-	4
Biology	5	1	-	-	-	6
Chemistry	1	-	-	-	-	1
Earth science	7	-	-	-	-	7
Ecology	3	-	1	-	-	4
Economics	1	-	-	-	-	1
Education	-	-	1	-	-	1
Engineer	4	2	-	-	-	6
Environmental science	-	1	1	-	-	2
Ethnic and cultural	9	-	-	1	-	10
General	22	10	11	2	3	48
Library science	2	1	-	-	-	3
Linguistic	1	-	-	-	-	1
Material sciences	1	-	-	-	-	1
Medecine and health	7	1	1	-	-	9
Public administration	3	-	-	-	-	3
Social science and humanity	7	2	-	1	-	10

Overall, this cross-analysis shows that all disciplines are covered by one or several repositories (except “Education”) and some disciplines may even need to converge in their efforts as multiple repositories are available in their area. While this shows interest and efforts of the community towards serving SAs, it may also risk confusing final users of the same domain (who may find similar resources in different non synchronised SACs). In addition, maintaining a single SAC per discipline may be easier to maintain than sustaining multiple ones.

Cross-analysis of disciplines 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 "Education" and "Environmental Science" from this analysis, since the technologies used for their SACs fall into the "Other" category (Fig. 6). Consequently, 15 out of the 17 disciplines are analysed.

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 over-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 a 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 15 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, 7 out of the 15 disciplines use a unique technology. “Economics”, "Linguistics" and “Public administration” use only Skosmos, "Library science" uses iQvo, "Chemistry" uses OLS, and “Agricultural” and "Material Science" use OntoPortal.

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

Generic technology / Record Count													
Standardized Wikipedia Disci...	OLS	ORR	OntoPortal	OpenTheso	Other	SKOSMOS	ShowVoc	TemaTres	Termweb	VocPrez	Vocab	iQvoc	Total ...
Agricultural science	-	-	1	-	2	-	-	-	-	-	-	-	3
Astronomy	-	-	1	-	3	-	-	-	-	-	-	-	4
Biology	2	-	1	-	3	-	-	-	-	-	-	-	6
Chemistry	1	-	-	-	-	-	-	-	-	-	-	-	1
Earth science	-	2	1	-	1	1	-	-	-	2	-	-	7
Ecology	-	-	2	-	1	-	1	-	-	-	-	-	4
Economics	-	-	-	-	-	1	-	-	-	-	-	-	1
Education	-	-	-	-	1	-	-	-	-	-	-	-	1
Engineer	1	-	1	-	3	1	-	-	-	-	-	-	6
Environmental science	-	-	-	-	2	-	-	-	-	-	-	-	2
Ethnic and cultural	-	-	-	2	7	-	-	1	-	-	-	-	10
General	1	-	-	-	33	9	1	1	1	1	1	-	48
Library science	-	-	-	-	2	-	-	-	-	-	-	1	3
Linguistic	-	-	-	-	-	1	-	-	-	-	-	-	1
Material sciences	-	-	1	-	-	-	-	-	-	-	-	-	1
Medecine and health	1	-	1	-	7	-	-	-	-	-	-	-	9
Public administration	-	-	-	-	1	2	-	-	-	-	-	-	3
Social science and humanity	-	-	-	2	3	5	-	-	-	-	-	-	10
Total général	6	2	9	4	69	20	2	2	1	3	1	1	120

FAIR-enabling assessments of the Semantic Artefact Catalogues

In this section, we extract some lessons learned from the assessment of the FAIR-enabling dimensions. The FAIR-enabling dimensions assessment per type of SACs is synthesised in Fig. 7.

Table 2. Number of type of SACs assessed with FAIR-enabling dimensions

	Assessed	Non assessed	Total
Active	66	54	120
Retired	3	58	61
Prototype	1	8	9

FAIR-enabling by type of SACs

As shown in Fig. 7, none of the FAIR-enabling dimensions are addressed by the “Listing” type of SAC. All dimensions were assessed to level 1 (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.

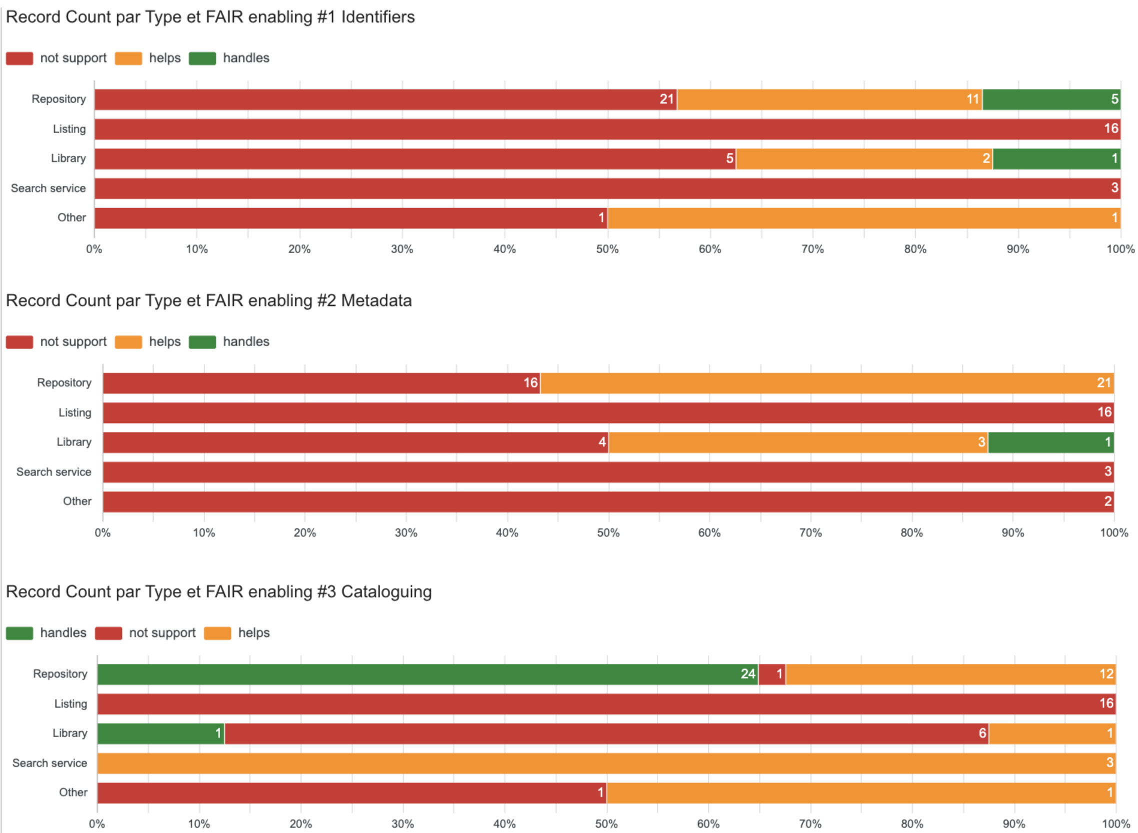
“Search service” and “Other” types of SACs do not support FAIR, with the exception of the criteria “Cataloguing” that was assessed to level 2 for most of them. Indeed, this acknowledges that when a semantic artefact is searchable by these SACs, it eases their findability. Still, we have not identified any active “Search service” that would completely address this cataloguing dimension (level 3) as it used to be done by retired semantic web search engines such as Swoogle [32] and Watson [33].

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 Vocabularies. These two 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 maintainers have set up the mechanisms to take care (level 3) 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.

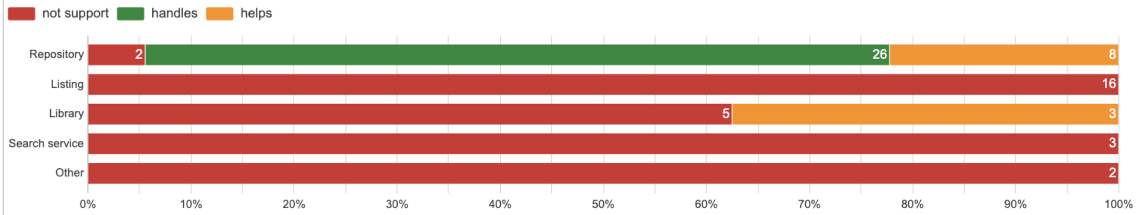
Finally, we have assessed 39 over 77 (51%) of the SAC of type "Repository". In this type of SAC, multiple FAIR-enabling dimensions are diversely addressed. Typically, the 3 dimensions Cataloguing, Resolvability, and Access Protocols are generally well addressed (often at level 3) by SACs, which sounds coherent as 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 3 where most of the SACs 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, SACs of type "Repository" really help/support, at least up to level 2, multiple FAIR-enabling dimensions; still some dimensions are mostly poorly addressed (level 1) such as Versioning, Reuse and mappings and Provenance.

This analysis shows SAC maintainers where to put some energy now in order to continue and enhance their support to SA developers in making their artefact FAIR.

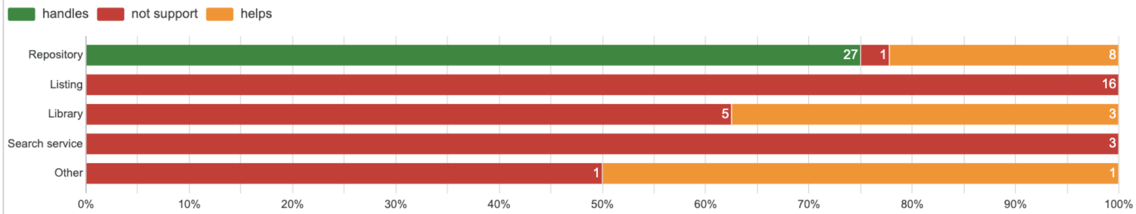
Figure 7. FAIR-enabling dimensions assessment per type of SACs



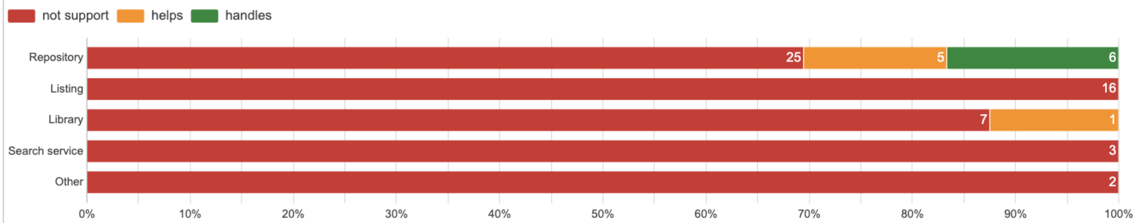
Record Count par Type et FAIR enabling #4 Resolvability

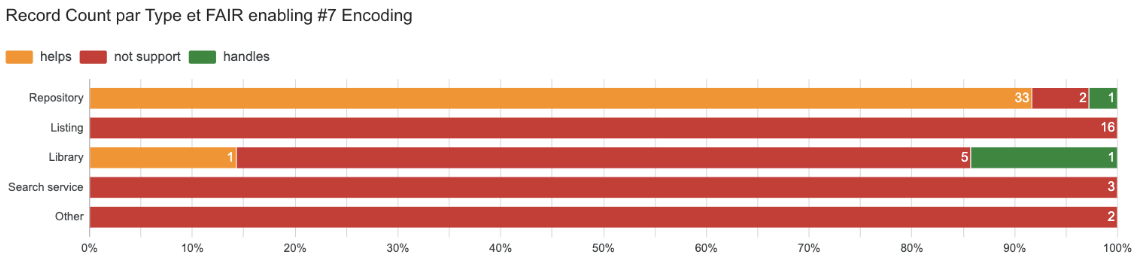


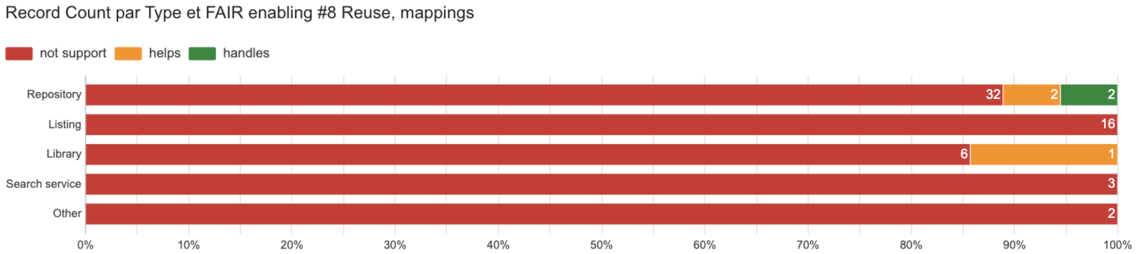
Record Count par Type et FAIR enabling #5 Access protocols

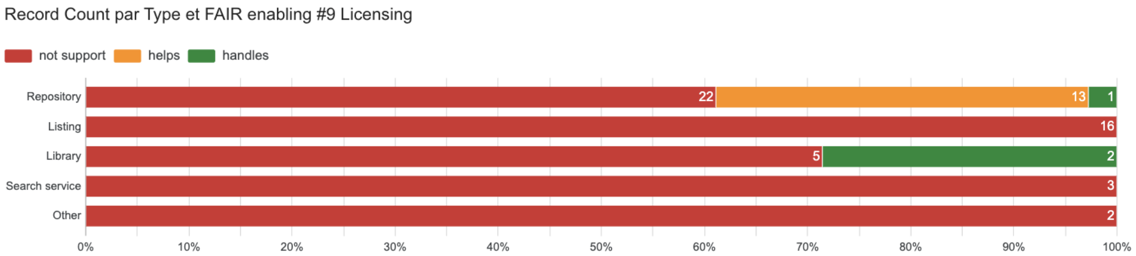


Record Count par Type et FAIR enabling #6 Versioning









FAIR-enabling by generic technology

Some technologies natively support some FAIR-enabling dimensions. Either the technologies were originally addressing the dimension or recent changes (since the publication and prominence of FAIR principles) were explicitly made to become more FAIR-enabling. For instance, the three dimensions Cataloguing, Resolvability, and Access protocols are well supported by "OntoPortal" and "OLS" or "VocPrez". The only technology where the Versioning dimension is always assessed level 3 or 2 is for "OntoPortal" which shows that this technology is the only one dealing with SA versioning criteria. Catalogues such as AgroPortal, EcoPortal and EarthPortal even explicitly declared a level 3 support for Versioning as these catalogues declare curating versioning information (dates, version info, new metadata, etc.) [9].

Similarly, some dimensions are not addressed 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 (e.g., 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.

Conclusions and discussion

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, as recognised in international efforts towards semantic interoperability (i.e., within 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.

This work provides the first large-scale, structured, and historical review of Semantic Artefact Catalogues, highlighting both the richness and the fragmentation of the ecosystem. By analysing 190 catalogues across disciplines, types, and technologies, we confirm that SACs are now essential components of the semantic infrastructure supporting FAIR data. The findings presented in this study provide a detailed overview of the current landscape of SACs. This diversity reflects the tailored approaches different communities have adopted to meet their specific needs. However, our results also reveal several important gaps and challenges that need to be addressed to improve SACs sustainability, interoperability, and overall impact.

First, the landscape remains highly fragmented. A large proportion of catalogues are domain-agnostic, while others coexist within the same discipline without clear coordination, leading to duplication of effort and potential confusion for users. This is further reinforced by the diversity of technologies employed, with more than half of the catalogues relying on ad-hoc solutions. A few generic technologies are well adopted, but their use remains uneven across disciplines, and convergence towards broader governance and interoperability is still needed. In addition, the significant number of retired catalogues highlights persistent sustainability issues and the difficulty of maintaining such infrastructures over time.

Second, our FAIR-enabling analysis shows that while SACs play a crucial role in supporting FAIR principles, their contribution is uneven across dimensions. Core technical aspects such as cataloguing, resolvability, and access protocols are generally well supported, particularly by repository-type SACs. In contrast, dimensions such as provenance, reuse, mappings, or rich metadata are less consistently addressed and often remain largely dependent on the efforts of semantic artefact providers rather than being enforced or facilitated by the catalogues themselves. This indicates that, beyond providing access, many SACs could further evolve towards more active support of their SAs developers' communities towards embracing FAIR and providing high quality SAs.

At the same time, the FAIR-enabling assessment presented in this work should be understood as an initial, coarse-grained analysis. It was intentionally designed to provide a broad comparative overview rather than a detailed benchmarking of individual catalogues. As such, the assessment relies partly on expert knowledge and interpretation, and it does not systematically capture all implementation nuances of each SAC. Future work could significantly improve this aspect by refining the evaluation framework, defining more precise and measurable criteria, and involving SAC providers in a more systematic validation process. In particular, aligning the assessment more closely with existing FAIR evaluation tools, and possibly operationalising the dimensions through automated or semi-automated indicators, would strengthen both the robustness and reproducibility of the analysis. Expanding the coverage to all identified SACs would also provide a more complete picture of the ecosystem.

Finally, an important outcome of this work is the establishment of the SAC review as a living, continuously updated resource. Between the initial version of the catalogue produced within the FAIR-IMPACT project (V1) and the current version presented in this paper (V2), we have already identified and recorded approximately 20 additional SACs. This rapid evolution of the landscape demonstrates both the dynamism of the field and the necessity of maintaining an up-to-date, community-driven overview. It also reinforces the importance of formally publishing this review: beyond reporting a static snapshot, the article provides a structured framework, a shared reference point, and an openly accessible dataset that can be extended and refined over time by the community.

In conclusion, this work not only documents the current state of Semantic Artefact Catalogues but also highlights key directions for future developments: fostering convergence towards shared and sustainable technologies, strengthening FAIR-enabling capabilities, and supporting coordinated community efforts such as the MOD-API support for SAC interoperability. We aim to contribute to a collective effort towards a more interoperable, robust, ecosystem for serving FAIR semantic artefacts.

Acknowledgments

This work has been supported by the European Project FAIR-IMPACT (101057344). We thank the members of the project that have provided feedback and/or references to SACs. We also thank the technology providers for their responses to our questions.

Author contribution

The contributions of the authors are described using the CRediT (Contributor Roles Taxonomy): CJ led the conceptualization of the study, project administration, and funding acquisition. CJ and NG developed the methodology, consolidated the data, analysed the results and wrote the original document. The other co-authors contribute to the review and FAIR-enabling assessment (validation and investigation and data curation). All authors contributed to the review and editing of the manuscript.

Appendices

Grouping of criteria, questions of recommendations

Here we regroup the criteria, questions or recommendations of the 5 state-of-the-art tools and methods for FAIR assessment of SAs. When a line is bolded, it recognises 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).

Legend:

- FxQx => a question from O'FAIRe [55];
- ACROx => a criterion from FOOPS! [56];
- P-Rec. x => a recommendation from FAIRsFAIR [1]
- Rule x => a recommendation from 10-simple rules [57]
- FVF-x => a criterion from FVF [58]

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. 1: Globally Unique, Persistent and Resolvable Identifiers must be used for Semantic Artefacts, their content (terms/ concepts/ classes and relations), and their version.
- P-Rec. 2: Globally Unique, Persistent, and Resolvable Identifiers must be used for Semantic Artefact 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.

Metadata

- F2Q1: Is the ontology described with additional 'MIRO must' metadata properties?
- OM1: This check verifies if 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
- 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. 3: A 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.

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

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. 9: Semantic 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.

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.**
● 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.

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. 8: Human 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.

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.

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.
- 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.

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?**

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
● 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.

References

1. Hugo, W., Le Franc, Y., Coen, G., Parland-von Essen, J., Bonino, L.: D2.5 FAIR Semantics Recommendations Second Iteration. (2020). https://doi.org/10.5281/ZENODO.5362010.

2. Corcho, O., Ekaputra, F.J., Heibi, I., Jonquet, C., Micsik, A., Peroni, S., Storti, E.: A maturity model for catalogues of semantic artefacts. Sci. Data. 11, 1–13 (2024). https://doi.org/10.1038/S41597-024-03185-4.

3. Nestola, E., Sgrigna, G., Ingrosso, G., Tarallo, A., Raho, D., Di Muri, C., Muresan, A.N., Rosati, I.: Assessing the adoption of the FAIR principles in Italian environmental research infrastructures. Patterns. 7, 101420 (2026). https://doi.org/10.1016/j.patter.2025.101420.

4. 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.

5. Di Muri, C., Pulieri, M., Raho, D., Muresan, A.N., Tarallo, A., Titocci, J., Nestola, E., Basset, A., Mazzoni, S., Rosati, I.: Assessing semantic interoperability in environmental sciences: variety of approaches and semantic artefacts. Sci. Data. 11, 1055 (2024). https://doi.org/10.1038/S41597-024-03669-3;SUBJMETA=172,496,648,704,706;KWRD=ENVIRONMENTAL+SCIENCES,RESEARCH+MANAGEMENT.

6. Jonquet, C.: Ontology Repository and Ontology-Based Services – Challenges, contributions and applications to biomedicine & agronomy, https://tel.archives-ouvertes.fr/, (2019).

7. Commission, E., for Research, D.-G., Innovation, Board, E.E., Corcho, O., Eriksson, M., Kurowski, K., Ojsteršek, M., Choirat, C., Sanden, M. v. d., Coppens, F.: EOSC interoperability framework – Report from the EOSC Executive Board Working Groups FAIR and Architecture. Publications Office (2021). https://doi.org/doi/10.2777/620649.

8. Jonquet, C., Grau, N., Poveda-Villalon, M., Garijo, D., Wilson, A., Rosati, I., Aubin, S., Corre, C., Alviset, G., Pulieri, M.: D4.2 - FAIR semantic artefact lifecycle from engineering, to sharing. Zenodo (2025). https://doi.org/10.5281/ZENODO.14643279.

9. Jonquet, C., Bouazzouni, S., Alviset, G., Fiore, N., Karam, N., Kihal, B., Pierkot, C., Pulieri, M., Rosati, I.: Federated FAIR Semantic Artefacts Discovery and Search with OntoPortal Federation. In: 24th International Semantic Web Conference. pp. 434–450. Springer, Nara, Japan (2025). https://doi.org/10.1007/978-3-032-09530-5_25.

10. Wilson, A., Jonquet, C.: D4.3 - Specification of shared metadata description of semantic artefacts and their catalogues including common reference API. Zenodo (2024). https://doi.org/10.5281/ZENODO.12579779.

11. 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.

12. 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).

13. Jonquet, C., Grau, N.: M4.4 - Review of Semantic Artefact Catalogues and guidelines for serving FAIR semantic artefacts in EOSC. Zenodo (2024). https://doi.org/10.5281/ZENODO.12799796.

14. 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).

15. Voß, J., Ledl, A., Balakrishnan, U.: Uniform description and access to Knowledge Organization Systems with BARTOC and JSKOS. Presented at the (2016). https://doi.org/10.7892/BORIS.143881.

16. Pesce, V., Tennison, J., Mey, L., Jonquet, C., Toulet, A., Aubin, S., Zervas, P., Pesce, V., Tennison, J., Mey, L., Jonquet, C., Toulet, A., Aubin, S., Zervas, P.: A Map of Agri-food Data Standards. (2018). https://doi.org/10.7490/F1000RESEARCH.1115260.1.

17. 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.

18. Sansone, S.A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A.L., Thurston, M.: FAIRsharing as a community approach to standards, repositories and policies. Nature Biotechnology 2019 37:4. 37, 358–367 (2019). https://doi.org/10.1038/s41587-019-0080-8.

19. Khan, Z., Keet, C.: The foundational ontology library ROMULUS. Model and Data Engineering. 200–211 (2013).

20. Gangemi, A.: Ontology Design Patterns for Semantic Web Content. In: Gil, Y., Motta, E., Benjamins, V.R., and Musen, M.A. (eds.) 4th International Semantic Web Conference, ISWC 2005. pp. 262–276. Springer, Galway, Ireland (2005). https://doi.org/10.1007/11574620_21.

21. 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.

22. 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.

23. Raphael Volz: ONTOSERVER —Infrastructure for the Semantic Web. In: First Semantic Web Working Symposium. pp. 96–98. , Stanford University, California, USA (2001).

24. Buitelaar, P., Eigner, T., Declerck, T.: OntoSelect: A Dynamic Ontology Library with Support for Ontology Selection. In: Demonstration Session at the 3rd International Semantic Web Conference, ISWC'04. pp. 3–6. , Hiroshima, Japan (2004).

25. Zhang, Y., Vasconcelos, W., Sleeman, D.: OntoSearch: An Ontology Search Engine. In: Bramer, M., Coenen, F., and Allen, T. (eds.) 24th SGAI International Conference on

Innovative Techniques and Applications of Artificial Intelligence, AI'04. pp. 58–69. Springer, Cambridge, UK (2004).

26. Z.Pan, J., Thomas, E.J., Sleeman, D.H., Pan, J.Z.: OntoSearch2: Searching and Querying Web Ontologies. In: IADIS International Conference WWW/Internet. pp. 211–219. , Murcia, Spain (2006).

27. D'Aquin, M., Lewen, H.: Cupboard - A place to expose your ontologies to applications and the community. In: 6th European Semantic Web Conference, ESWC'09. pp. 913–918. Springer, Heraklion, Crete, Greece (2009). https://doi.org/10.1007/978-3-642-02121-3_81.

28. 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).

29. Obrst, L., Baclawski, K.: The OOR Initiative – An Update Announcing the new OOR Logo Designed by : Ali Hashemi. 2011, 1–23 (2011).

30. Jonquet, C., Graybeal, J., Bouazzouni, S., Dorf, M., Fiore, N., Kechagioglou, X., Redmond, T., Rosati, I., Skrenchuk, A., Vendetti, J.L., Musen, M.: Ontology Repositories and Semantic Artefact Catalogues with the OntoPortal Technology. In: Semantic Web, ISWC 2023. ISSN : 0302-9743. pp. 38–58. Springer, Athens, Greece (2023). https://doi.org/10.1007/978-3-031-47243-5_3.

31. 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).

32. Finin, T., Ding, L., Pan, R., Joshi, A., Kolari, P., Java, A., Peng, Y.: Swoogle : Searching for Knowledge on the Semantic Web ∗. Data Engineering. 1682–1683 (2004).

33. 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).

34. 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).

35. Osma Suominen, Sini Pessala, Jouni Tuominen, Mikko Lappalainen, Susanna Nykyri, Henri Ylikotila, Matias Frosterus, Eero Hyvönen: Deploying national ontology services: From ONKI to Finto. In: International Semantic Web Conference. CEUR, Riva del Garda, Italy, (2014).

36. Xiang, Z., Mungall, C., Ruttenberg, A., He, Y.: Ontobee: A Linked Data Server and Browser for Ontology Terms. In: Bodenreider, O., Martone, M.E., and Ruttenberg, A. (eds.) 2nd

International Conference on Biomedical Ontology, ICBO’11. p. 3. , Buffalo, NY, USA (2011).

37. 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.

38. Slater, L., Gkoutos, G. V, Schofield, P.N., Hoehndorf, R.: Using AberOWL for fast and scalable reasoning over BioPortal ontologies. Biomedical Semantics. 7, 49 (2016). https://doi.org/10.1186/s13326-016-0090-0.

39. Grosjean, J., Soualmia, L.F., Bouarech, K., Darmoni, S.J., Jonquet, C.: Comparing BioPortal and HeTOP: towards a unique biomedical ontology portal? In: (2nd International Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO’14. p. 11. , Granada, Spain (2014).

40. 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.

41. 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).

42. Science, D., Hahnel, M., Smith, G., schoenenberger, henning, Scaplehorn, N., Day, L.: The State of Open Data 2023. (2023). https://doi.org/10.6084/M9.FIGSHARE.24428194.V1.

43. 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).

44. Jonquet, C., Emonet, V., Musen, M.A.: Roadmap for a multilingual BioPortal. In: Gracia, J., McCrae, J.P., and Vulcu, G. (eds.) 4th Workshop on the Multilingual Semantic Web, MSW4’15. pp. 15–26. , Portoroz, Slovenia (2015).

45. 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.

46. Kokkinaki, A., Luong, Q., Thompson, C., Car, N., Moncoiffe, G., Kokkinaki, A., Luong, Q., Thompson, C., Car, N., Moncoiffe, G., Kokkinaki, Alexandra, Luong, Quyen, Thompson, Christopher, Car, Nicholas, Moncoiffe, Gwenaelle: Applying VocPrez to operational semantic repositories: the NVS experience. In: EGUGA. pp. EGU21-15669 (2021). https://doi.org/10.5194/EGUSPHERE-EGU21-15669.

47. Stellato, A., Fiorelli, M., Lorenzetti, T., Turbati, A., van Gemert, W., Dechandon, D., Gerencsér, A., Francesconi, E.: ShowVoc: A Thorough Platform for Publishing and

Browsing Linked Open Datasets. In: Extended Semantic Web Conference. pp. 316–333. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-94578-6_18.

48. Stellato, A., Rajbhandari, S., Turbati, A., Fiorelli, M., Caracciolo, C., Lorenzetti, T., Keizer, J., Pazienza, M.T.: VocBench: A web application for collaborative development of multilingual thesauri. In: LNCS. pp. 38–53. Springer Verlag (2015). https://doi.org/10.1007/978-3-319-18818-8_3.

49. Fiore, N., Magagna, B., Goldfarb, D.: EcoPortal: A Proposition for a Semantic Repository Dedicated to Ecology and Biodiversity. In: A. Algergawy, Karam, N., Klan, F., and Jonquet, C. (eds.) 2nd International Workshop on Semantics for Biodiversity, S4BioDiv'17. CEUR, Vienna, Austria (2017).

50. Pierkot, C., Alviset, G., Vernet, M.: The EarthPortal, towards an ontology repository for the Earth System semantic artefacts. In: 3rd Workshop on Ontologies for FAIR and FAIR Ontologies (Onto4FAIR). , Leipzig, Germany (2023). https://doi.org/10.34894/VQ1DJA.

51. Karam, N., Fillies, J., Jonquet, C., Syphax Bouazzouni, ·, Löffler, F., Zander, F., König-Ries, B., Güntsch, A., Diepenbroek, M., Paschke, A.: BiodivPortal: Enabling Semantic Services for Biodiversity within the German National Research Data Infrastructure. Datenbank-Spektrum 2024 24:2. 24, 129–137 (2024). https://doi.org/10.1007/S13222-024-00474-5.

52. Cecconi, B., Debisschop, L., Derrière, S., Louys, M., Corre, C., Grau, N., Jonquet, C.: OntoPortal-Astro, a semantic artefact catalogue for astronomy. Astronomy and Computing. 53, 100991 (2025). https://doi.org/10.1016/J.ASCOM.2025.100991.

53. Nyberg Åkerström, W., Baumann, K., Corcho, O., David, R., Le Franc, Y., Madon, B., Magagna, B., Micsik, A., Molinaro, M., Ojsteršek, M., Peroni, S., Scharnhorst, A., Vogt, L., Widmann, H.: Developing and implementing the semantic interoperability recommendations of the EOSC Interoperability Framework. (2024). https://doi.org/10.5281/ZENODO.10843882.

54. Gonzales-Aguilar, A., Ramírez-Posada, M., Ferreyra, D.: TemaTres: software para gestionar tesauros. El Profesional de la Informacion. 21, 319–326 (2012).

55. 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.

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

57. Malone, J., Stevens, R., Jupp, S., Hancocks, T., Parkinson, H., Brooksbank, C.: Ten Simple Rules for Selecting a Bio-ontology. PLoS Comput. Biol. 12, 6 (2016). https://doi.org/10.1371/journal.pcbi.1004743.

58. Xu, F., Juty, N., Goble, C., Jupp, S., Parkinson, H., Courtot, M.: Features of a FAIR vocabulary. Journal of Biomedical Semantics 2023 14:1. 14, 6- (2023). https://doi.org/10.1186/S13326-023-00286-8.

59. Matentzoglu, N., Malone, J., Mungall, C., Stevens, R.: MIRO: guidelines for minimum information for the reporting of an ontology. J. Biomed. Semantics. 9, 6 (2018). https://doi.org/10.1186/s13326-017-0172-7.

60. Rigneault, H., Kumar, N.G., Cossart, R., Septier, D., Brévalle-Waslilewki, G., Kudlinski, A., Kaszas, A.: Are those URIs so cool? URI Resolution and Content Negotiation in Ontology Repositories. In: 24th International Semantic Web Conference, Demo and Poster. p. 255. CEUR (2025). https://doi.org/10.34894/VQ1DJA.