



OntoPortal Workshop 2024 Report

Clement Jonquet, Jennifer L Vendetti, Guillaume Alviset, Ilaria Rosati, John Graybeal, Naouel Karam, Mark Musen, Michael Dorf, Alex Skrenchuk, Syphax Bouazzouni, et al.

► To cite this version:

Clement Jonquet, Jennifer L Vendetti, Guillaume Alviset, Ilaria Rosati, John Graybeal, et al.. OntoPortal Workshop 2024 Report. INRAE; Stanford University. 2024. hal-04891214

HAL Id: hal-04891214

<https://hal.science/hal-04891214v1>

Submitted on 16 Jan 2025

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

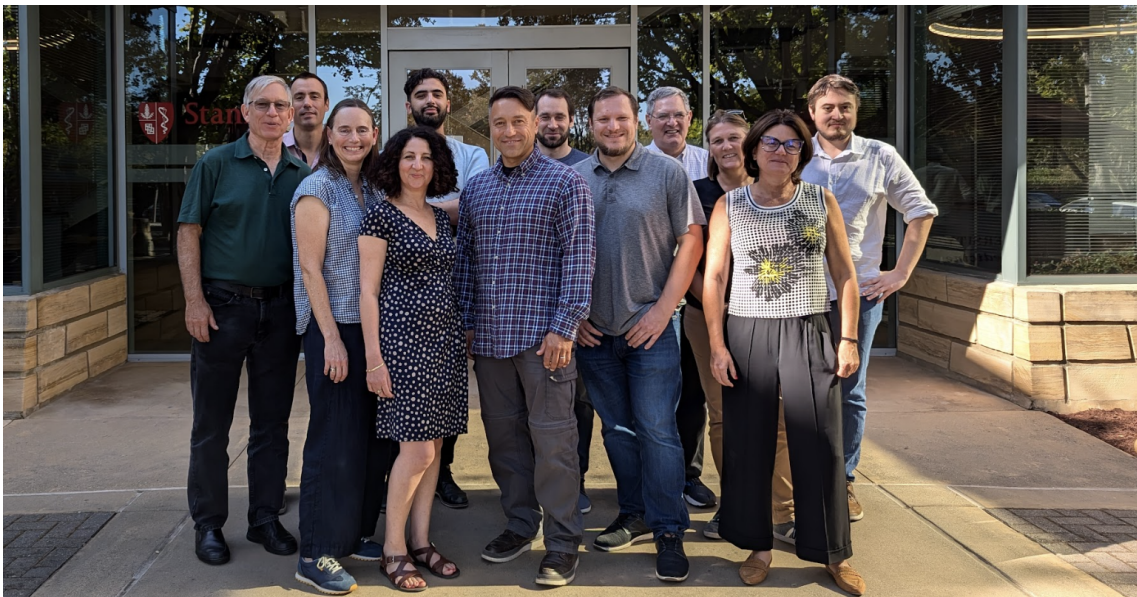
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



OntoPortal Workshop 2024 Report

Hosted by Stanford-BMIR,
Stanford, USA, September 23-26th, 2024

Clement Jonquet, Jennifer Vendetti, Guillaume Alviset, Ilaria Rosati, John Graybeal, Naouel Karam, Mark Musen, Michael Dorf, Alex Skrenchuk, Syphax Bouazzouni, Christelle Pierkot,
and the members of the OntoPortal Alliance



Participants: Guillaume Alviset, Clement Jonquet, Ilaria Rosati, John Graybeal, Naouel Karam, Mark Musen, Michael Dorf, Alex Skrenchuk, Syphax Bouazzouni, Christelle Pierkot, Jennifer Vendetti

Online participants: Martina Puileri, Bilel Kihal, Alexandru-Aurelian Todor, Jan Fillies

Other guests: Baptiste Cecconi, Beatrice Markhoff, Yuriy Marykovskiy, Matthew Horridge, Christian Kindermann

Executive Summary

The [OntoPortal Alliance](#) is a consortium of several research and infrastructure teams dedicated to promoting the development of ontology repositories—in science and other disciplines—based on the open, collaboratively developed [OntoPortal open source software](#). Teams in the Alliance develop and maintain several openly accessible semantic artefact catalogues or ontology repositories in multiple disciplines. The list of repositories is available on the Alliance [web page](#) and the [teams are listed on GitHub](#).

The OntoPortal Alliance members and other partners gathered from Sept. 23 to 26th, 2024 in Stanford for the 3rd OntoPortal Workshop. The meeting gathered around 20 persons including management, research and technical profiles (13 on site and 7 remotely). The 2024 workshop's main goals were to consolidate the OntoPortal Alliance organization and shared agenda, and to engage with newcomers. The program included several sessions (technical, content, management) as well as a newcomers session which featured three new use cases in astronomy, cultural heritage and for general semantic web vocabularies (LOV).

The workshop allowed the Alliance to re-align on the questions related to the management of the upstream OntoPortal code repository (<https://github.com/ontportal>), as well as the distribution of the OntoPortal virtual appliance. Multiple new interfaces and new features were demonstrated such as URI management or new search or SPARQL query capabilities. The workshop also demoed the first federation of four OntoPortal-based repositories with the upcoming federation joint release of Agro/Eco/Earth/BiodivPortal. Different prototypes were presented including revised work related to the Annotator using LLMs, a (BioPortal-based) CustomGPT.

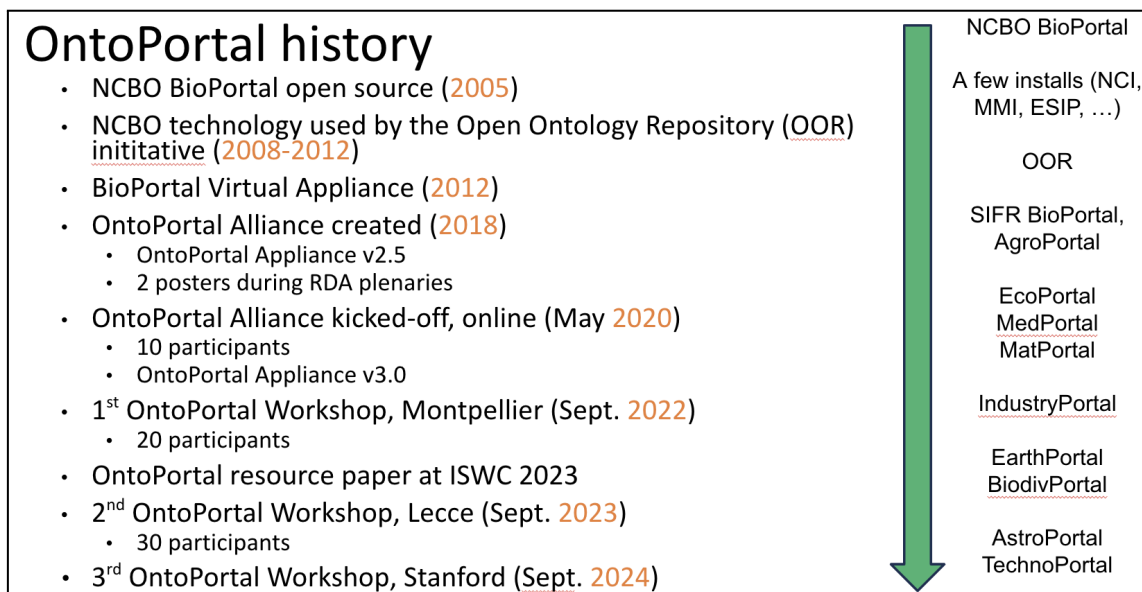
The event demonstrated the clear motivation and commitment of the Alliance members to collaborate on the OntoPortal technology and the implementation of ontology-based services in general. This document is a summary of the discussions and the decisions taken.

OntoPortal year in review

This workshop was the third physical workshop of a series kicked-off in 2020. The group started to work on the outcomes of previous workshops reported here:

- OntoPortal Workshop 2022 Report: <https://hal.science/hal-04087929>
- OntoPortal Workshop 2023 Report: <https://hal.science/hal-04494573>

An updated review of the things discussed and planned during previous workshops is available later in this report.



The recent achievements of the last 12-18 months were summarized and recalled to the whole group. Including:

- The publication of a reference paper at ISWC 2023:
https://dx.doi.org/10.1007/978-3-031-47243-5_3
- The publication of several OntoPortal related research papers (see next section)
- The release of a new website at: <https://ontoportal.org>
- The production of an updated architecture diagram (see next section)
- The arrival of two new public portals released or about to be publicly released
 - OntoPortal-Astro in the domain of astronomy and supported by Observatoire de Paris,
 - TechnoPortal (<https://technoportal.hevs.ch>) in the domain of technology sciences, supported by Ost / Hevs.
- The release of a shared documentation at: <https://ontoportal.org/documentation>
- The preparation of joint release with a smaller group (Agro/Eco/Earth/BiodivPortal) that will feature a first round of federated services.

The group analysed the meeting activities and organization of the past year (10 management meetings (average of 8-9 attendees) and 9 developers meetings (average of 5-6 attendees)) and assessed the model was working well to maximize information exchange while minimizing the work required. Still, we observed that besides the motivation of each group to lead its own portal/project, the engagement of each partner in and for the Alliance could be reinforced especially taking on tasks between meetings.

The group renewed its commitment to work as an open source project working with and maintaining the upstream code on GitHub (<https://github.com/ontoport>). Most of the projects are [now forks of the upstream code](#) (except MedPortal and MatPortal) which set the environment for sharing code. However, the group acknowledged we have not yet found the ideal model to do work with and maintain the upstream code, especially, handling the technical contributions proposed by the two main development teams at Stanford and Montpellier. Pre-workshop, only a few syncs were made by the BioPortal team and the AgroPortal team would refrain from making contributions to the upstream resulting in an undesirable situation where the upstream code doesn't correspond to any code "running" in any OntoPortal instance, nor was it used to compile the virtual appliance. The distance between the code on the *ncbo* repository and *ontoport-lirmm* repository was also considered too great, requiring a renewed strategy to realign the code. Also, we observed that discussion on GitHub –a mechanism previously identified as a solution to exchange views, track opinions and decisions– is still not really used.

The two core "actions" of the Alliance were discussed and updated:

- A. Provide a robust, stable and easy to set up OntoPortal appliance to projects or institutions interested in setting up their own, generally private, instance, but without requiring customization & development, and nor necessarily joining the Alliance.
- B. Respond and provide support to new use cases (project or community) who are setting up a, generally public, "reference" ontology repository, interested in doing so with the OntoPortal technology, and eventually modifying/contributing to the code and interested in joining the Alliance.

We observed historical action A is mostly addressed by the BioPortal team especially as the provider of the OntoPortal virtual appliance and the more recent action B is often handled by the AgroPortal team, often coming from research or infrastructure groups in Europe. The group recognized that the Alliance would need more communication and outreach to promote its missions and these two actions.

OntoPortal related papers

Jonquet, C., Graybeal, J., Bouazzouni, S., Dorf, M., Fiore, N., Kechagioglou, X., ... & OntoPortal Alliance. (2023, October). **Ontology repositories and semantic artefact catalogues with the ontoPortal technology**. In *22nd International Semantic Web Conference* (pp. 38-58), Athens, Greece. https://dx.doi.org/10.1007/978-3-031-47243-5_3

Karam, N., Fillies, J., Jonquet, C., Bouazzouni, S., Löffler, F., Zander, F., ... & Paschke, A. (2024). **BiodivPortal: Enabling Semantic Services for Biodiversity within the German National Research Data Infrastructure**. *Datenbank-Spektrum*, 1-9. <https://doi.org/10.1007/s13222-024-00474-5>

Corcho, O., Ekaputra, F. J., Heibi, I., Jonquet, C., Micsik, A., Peroni, S., & Storti, E. (2024). **A maturity model for catalogues of semantic artefacts**. *Scientific Data*, 11(1), 479. <https://doi.org/10.1038/s41597-024-03185-4>

Amdouni, E., Sarkar, A., Jonquet, C., & Karray, M. H. (2023, November). **IndustryPortal: a common repository for FAIR ontologies in industry 4.0**. In *22nd International Semantic Web Conference Demo & Poster session*. Athens, Greece. <https://hal.science/hal-04207343v1>

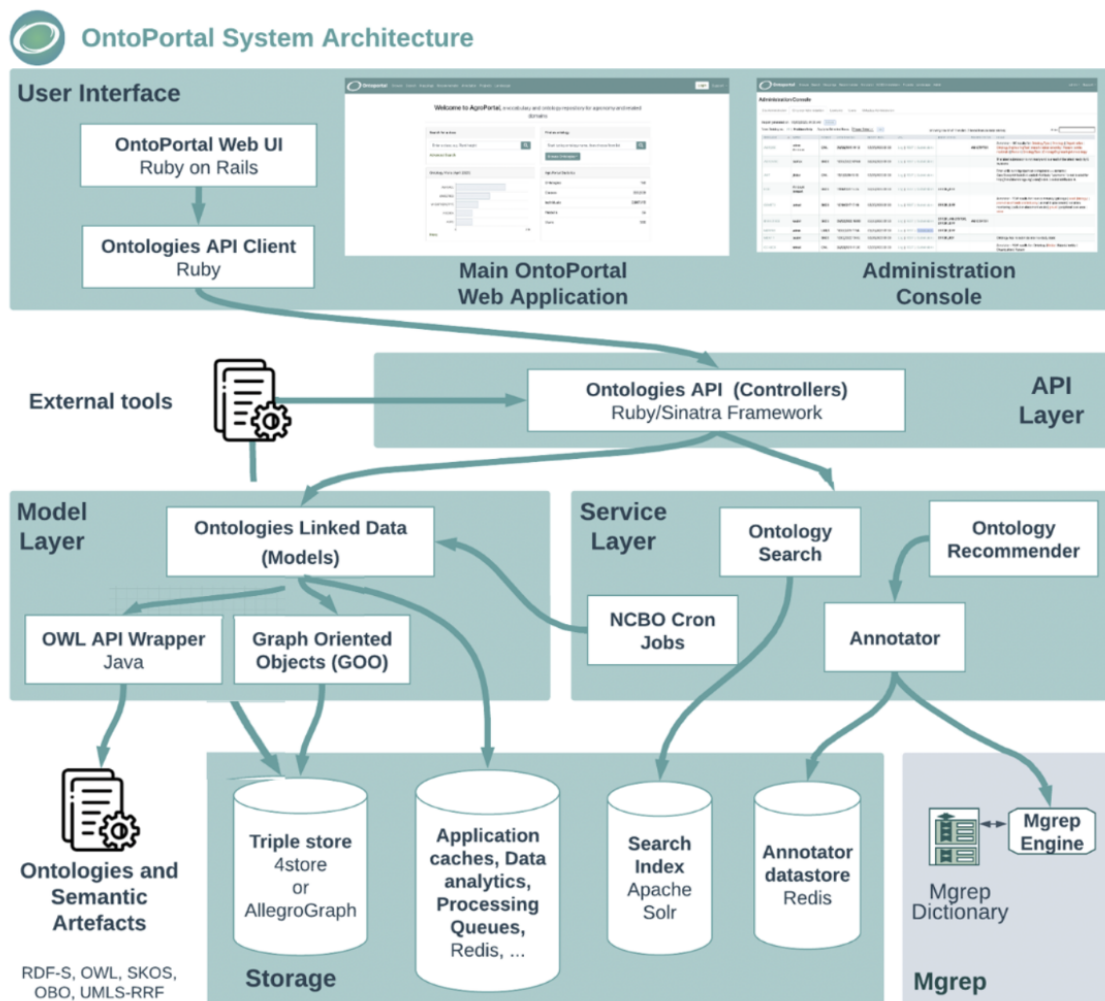
Alviset, G., Pierkot, C., & Vernet, M. (2023, September). **The EarthPortal, towards an ontology repository for the Earth System semantic artefacts**. In *3rd Workshop on Ontologies for FAIR and FAIR Ontologies (Onto4FAIR)* (p. 17), Leipzig, Germany. <https://hal.science/hal-04312604>

Rovetto, R. J. (2023). **AstroPortal: An ontology repository concept for astronomy, astronautics and other space topics**. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2309.10288> .

Tarallo, A., Pulieri, M., Ramezani, P., & Rosati, I. (2024). **Advancements in EcoPortal: Enhancing functionalities for the ecological domain semantic artefacts repository**. *FAIR Connect*, 2(1), 1-7. <http://dx.doi.org/10.3233/FC-240002>

Di Muri, C., Pulieri, M., Raho, D., Muresan, A. N., Tarallo, A., Titocci, J., ... & Rosati, I. (2024). **Assessing semantic interoperability in environmental sciences: variety of approaches and semantic artefacts**. *Scientific Data*, 11(1), 1-13. <https://doi.org/10.1038/s41597-024-03669-3>

OntoPortal updated architecture diagram

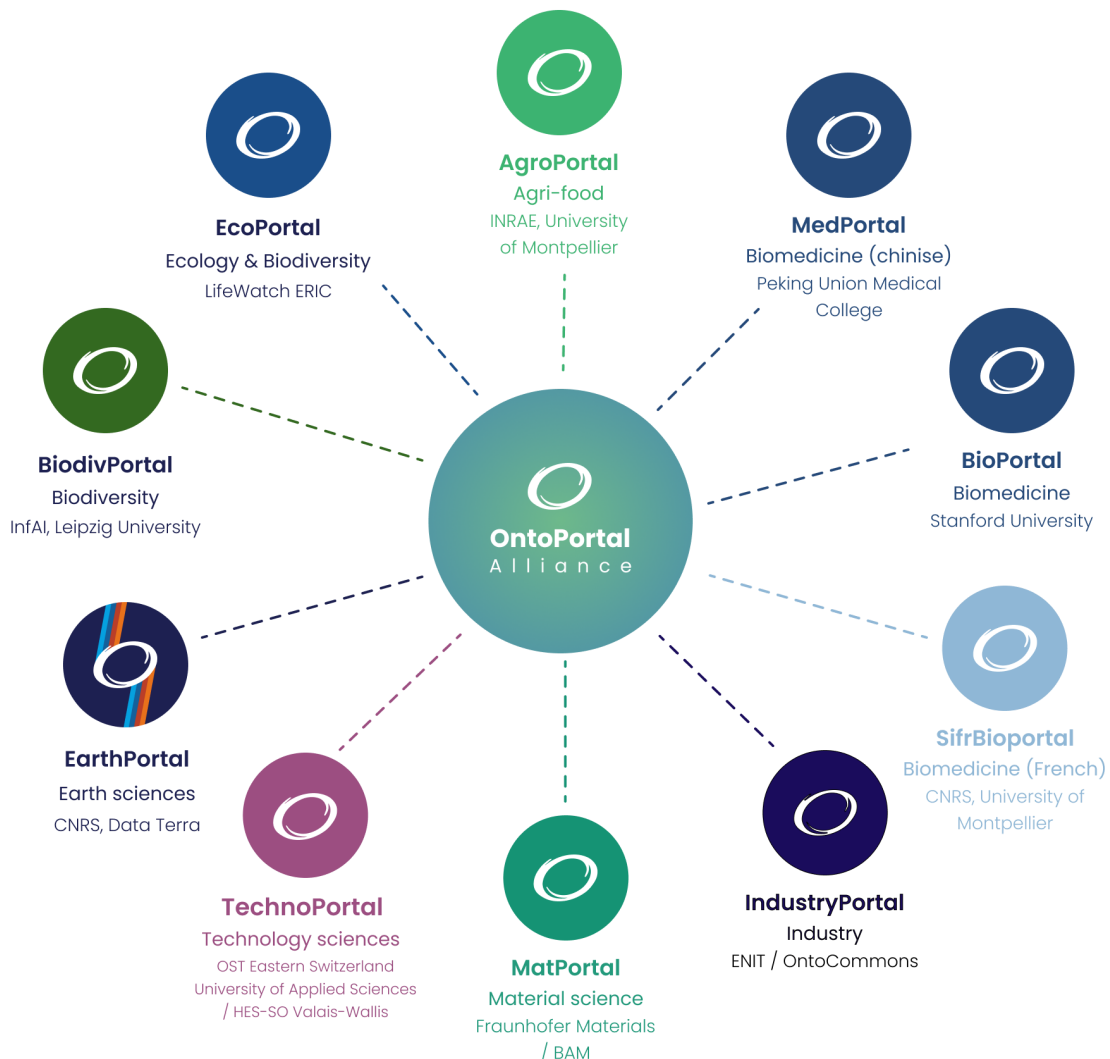


OntoPortal updated philosophy and statement

- OntoPortal's motivation is to **ease the use of ontologies and semantic artefacts in general**, especially for users not familiar with semantic technologies. OntoPortal becomes useful in the life cycle of semantic artefacts at the point when releasing a specific version to archive or share becomes necessary.
- OntoPortal installations generally **do not own ontologies or artefacts**, they just serve them. Most portals have adopted an open approach with a flexible editorial policy, where users can upload resources themselves, and the evaluation is left to the community. Generally, the OntoPortal project teams do not label an ontology as 'good' or 'bad' but provide the means (metadata, FAIRness assessment, use information, updates, etc.) so that final users can decide what ontology or artefact to use.

- Each portal has the **autonomy to establish its own editorial policy and determine the content it hosts**. This may range from supporting a limited and specific set of ontologies tailored to a particular group to hosting a comprehensive archive of all ontologies and artefacts within a domain. However, each portal is **accountable for the information and metadata it presents** about an ontology, including the curation procedures applied to ensure its accuracy and quality.
- It is accepted that semantic resources **can be hosted in several OntoPortal installations**, but better coordination of the content is needed. An approach to federating our content and services is necessary typically involving a canonical portal for duplicate ontologies that would serve the data (source file) and metadata to other portals.
- The OntoPortal Alliance does not yet have a **legal entity but will engage in the process of creating one** to ease the governance and decision making.
- OntoPortal **technology is as modular as we can make it** and contains the basic components that everyone wants. Any new functional feature developed by OntoPortal developers should be **designed as modular as possible with and can be parameterized** in the configuration files to be activated (or not) or tuned by each portal.
- We adopt an **open source philosophy** reflecting our belief that emphasizing shared code and development will maximize our individual progress. Since 2022 and going forward, we commit to continually increase our reliance on, and support for, the **primary GitHub repository** (<https://github.com/ontoportals>) which hosts the upstream code.
- The upstream OntoPortal code may be contributed by all (as pull requests) within or outside the Alliance. The Alliance technical teams engage in being sure this code is coherent, stable and usable (even if not directly used by the OntoPortal appliance) and represent the consensus of what the Alliance thinks relevant in terms of features and functionalities for OntoPortal.
- We rely on **standardized semantic web formats** and technologies to develop generic, domain agnostic, and compliant software that can be applied to a wide range of uses that may not be anticipated. Community or domain specific customization, if any, are developed within each specific OntoPortal instance.

Team presentations



BioPortal

Over the past year, we have made significant updates to the BioPortal technology stack, focusing on enhancing user experience, improving system performance, and expanding platform capabilities. The user interface has been modernized through an upgrade to Bootstrap 5.2.3, moving from the older version 4.1.0, and a transition to Rails 7.0.8, which together ensure greater security, efficiency, and access to more recent web technologies. Additionally, the account management pages have been redesigned to improve navigation and provide a more intuitive user experience. On the infrastructure side, we successfully migrated from CentOS 7 to Ubuntu 22.04 and

transitioned from 4store to AllegroGraph, resulting in enhanced system stability and performance.

We have also adopted basic multilingual support, allowing users to view ontology terms in their respective languages for ontologies defined in different languages. The search functionality has been improved with support for OBO IDs, and we resolved a bug that previously hindered short ID searches across many ontologies. To streamline access to the latest ontology content, we added a new API feature for on-demand remote ontology pulls. Furthermore, we addressed several issues with our search API and merged pull requests that improve multilingual support and fix earlier inconsistencies caused by the introduction of this feature.

These improvements reflect our ongoing commitment to providing a robust, efficient, and user-friendly platform for the BioPortal community.

AgroPortal

The past year has been highly productive for the AgroPortal team, which has grown in maturity and expertise. AgroPortal status update presented related work done in the context of FAIR-IMPACT (described after) and gave a tour of the technical releases achieved in the last 12 months. Since January 2024, we have implemented a dynamic workflow with monthly release cycles, reflecting our commitment to continuous improvement.

Details are available in [AgroPortal public release notes](#) with releases between v2.7.x and v3.0. A focus was given on work related to: (i) new indexing and search boxes; (ii) an embedded SPARQL query editor; (iii) URI resolution and negotiation service; (iv) optimization and (v) enhanced property support. The work done in the context of the federation was presented later in the workshop.

Releases 2.7.X - New UI and Multilingual 🌟 – From Dec. 2023 to June 2024

- Terms & conditions
- API call buttons
- Ontology selection component
- OntoPortals logos
- Agents
- Metrics and statistics page
- Export metadata to multiple formats
- Page load & cache optimization
- Multilingual support
- Multilingual or language-specific search
- Internationalized user interfaces
- Simplified ontology submission process
- Single Sign-on
- Admin analytics and monitoring
- Admin pages for groups and categories

Releases 2.8.X – Index & search everything 🐞 – From June 2024 to July 2024

- New search box
- Search anywhere
- SPARQL endpoint
- Reused objects icons
- Slices back

Releases 2.9.X – URI management 🔗 – From July 2024 to September 2024

- URI resolvability checker
- URI management (resolution & negotiation)
- Tool box (content converter, search by URI)

Releases 2.10.X – From September 2024

- Properties with properties
- Group and categories page

Releases 3.X.X – Federation and Virtuoso 🍷 – End of the year

AgroPortal, now nine years old, insisted on the fact that none of these developments and improvements were agri-food-specific developments and the platform remains generic and primed for broader reuse. The team's primary focus has shifted towards

advancing OntoPortal as a foundational framework, subsequently applying its innovations to AgroPortal and other OntoPortals we explicitly collaborate with, such as EcoPortal, EarthPortal, BiodivPortal, and of course BioPortal. Soon OntoPortal-Astro and LovPortal.

Among the elements in the upcoming year roadmap, the AgroPortal team mentioned: more consolidation of the software overall (more technical issues are currently being opened than closed), triple store migration, make the URI management service adopted, finish remaining new UIs (projects, landscape, notes), rebuild the Projects section including with integration with external databases, show logical axioms, social features (with harmonization of Notes, Reviews, Change-requests, Provisional objects and Notifications), implementation of MOD-API, new metadata curation campaign, upload half of those 200 identified candidates, activate validators and rules for metadata properties, support SKOS collections as primary vocabulary object, learn to deploy out the doors (EOSC services).

EcoPortal

This year, the EcoPortal team has been actively engaged in several key initiatives. One major focus was the testing of the new version of EcoPortal, which introduces several enhancements. Key updates include multilingual support, a redesigned user interface and user experience, a new tab-based metadata editing page, and an updated summary page. These features, along with the new documentation, will be officially released in December. Additionally, a patch release was issued in June, featuring updates to VocBench and ShowVoc. Another important activity was the restructuring of the documentation, carried out in collaboration with the Alliance. This partnership aimed to generalize and release a federated documentation system, enhancing accessibility and usability across platforms.

Several collaborations within national and international projects have been carried out during this year. As part of the Italian ITINERIS project, we conducted an extensive [analysis of approximately 500 semantic artefacts](#). This analysis evaluated their scientific domain coverage, management across various catalogues, and key aspects of their FAIRness. Among these artefacts, we are identifying and selecting those relevant to biodiversity and ecosystem communities that are not yet available in repositories. These artefacts will soon be uploaded to EcoPortal, further enriching its content. Within the ITINERIS framework, the Semantic Interoperability Profile (SIP) was compiled for EcoPortal with the objective of conducting a FAIR assessment of the FAIR supporting resources adopted within the repository to support the publication and maintenance of semantic artefacts. This work [resulted in a paper](#) detailing the new developments in version 2.0.0.

As part of the FAIR-IMPACT project, we proposed the integration of the LifeWatch Italy Data Portal and Metadata Catalogue with EcoPortal as a use case. This activity aims to improve the FAIRness of (meta)data produced within the biodiversity and ecosystem community through the use of semantic artefacts.

Lastly, we are working on updating EcoPortal's governance model, building on the progress made in Task 4.1 of the FAIR-IMPACT project. This ongoing effort aims to ensure robust and sustainable governance for the platform moving forward.

EarthPortal

Since EarthPortal went online shortly before the previous OntoPortal workshop at Lecce in 2023, endeavours were made toward gathering Data Terra clusters semantic artefacts and the connection with [the EaSy Data repository](#). Content-wise, 56 ontologies/vocabularies are already stored and 50 more are candidates to be and with a total of 25 registered users at the time of this workshop. The presentation of the EarthPortal to the annual Data Terra seminar sparked interest from the clusters, which will help to improve their artefacts with more and better metadata but also to suggest new use cases of EarthPortal tools with external applications. The connection from EaSy Data to the EarthPortal to enhance the deposit of datasets has already been implemented in the context of FAIR-IMPACT and is one of the examples of such use cases. Other collaborations are on the way, such as the use case done between FAIR-IMPACT and FAIRE-EASE projects that produced the Semantic analyzer tool to annotate metadata, using the EarthPortal within their knowledge base. There is also an ongoing collaboration with NFDI4Earth and EPOS MSL.

During the upcoming year, we plan to pursue existing collaborations and have new ones, to curate and manage the growing amount of resources and to enhance the technology. EarthPortal is also part of the ongoing efforts towards the federation of content with AgroPortal, EcoPortal and BiodivPortal and we also plan to also refresh and enhance the Project feature that has been lagging behind with the new releases.

BiodivPortal

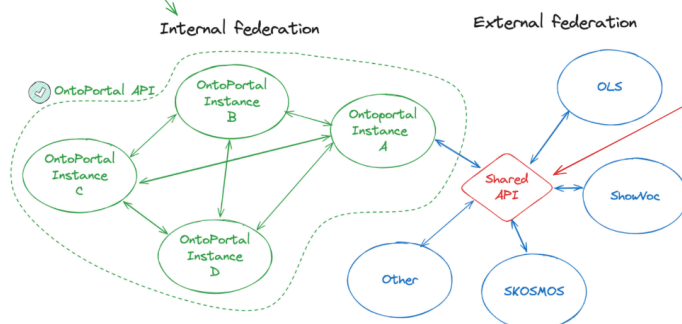
BiodivPortal has been successfully aligned with the latest AgroPortal release before the workshop. A significant milestone was achieved with the addition of 23 new ontologies, increasing the total to 46. The team started prototyping a new diff user interface (see image below) and accompanying tools, designed to provide an intuitive and detailed visualization of ontological changes over time. An additional output format is based on [KGCL](#) (Knowledge Graph Change Language) and will enable users to track modifications both at the ontology and individual term levels, ensuring enhanced transparency and usability for managing evolving ontologies.

Diff



Additionally, the team has begun developing an LLM Extension to the Annotator aimed at improving the annotation as well as detecting unknown concepts within texts, enhancing the Annotator capabilities and connecting to the term suggestion workflows. As part of the TS4NDFI project, efforts are underway to build a federated API Gateway. This gateway is designed to extend the federation work achieved by the OntoPortal Alliance on the user interface side, while introducing federation at the API level. Such enhancements will not only streamline interoperability with other technologies but also empower users with more efficient access to federated content, ultimately advancing the overall utility and scope of OntoPortal.

OntoPortal work



TS for NDFI work

Furthermore, we [published a paper entitled](#) *BiodivPortal: Enabling Semantic Services for Biodiversity within the German National Research Data Infrastructure*, which highlights the platform's role in supporting semantic technologies and services for biodiversity research.

Other news from groups non present

MatPortal, IndustryPortal and MedPortal did not present any status update.

The French **SIFR BioPortal** was not (re)presented either and C. Jonquet announced the portal will soon be discontinued as his research group does not work on it anymore and the need of a national/French language focus ontology/terminology service for the biomedical domain is now partly addressed by other/new institutional solution such as the French ANS's SMT (<https://smt.esante.gouv.fr>).

A. Todor presented some work related to the use of generative AI tools with **MatPortal** (cf. specific session).

X. Yang shared some news offline with the Alliance related to **MedPortal**. The focus of MedPortal's work is to promote its application in ontology construction and scientific data governance. Progress was made in two areas in 2024. First, a Python-based ontology construction tool named Py2Onto (<https://github.com/MedportalProject/Py2ONTO-Edit>): (i) automatic generation of ontologies from table templates; (ii) terminology mapping; (iii) ontology segmentation; (iv) translation of ontology terms using large language models. Second, the application of ontologies in biomedical data governance was advanced. In the genomic data management system of China National Population Health Data Center, it has helped achieve the description of metadata about investigation projects, experiments, and samples using ontology terms for genomic data submission. The system will be online soon.

With respect to the status of **IndustryPortal**, H. Karray and A. Sarkar have reported not working anymore at ENIT and not working anymore on IndustryPortal. T. Lounge has indicated being committed to maintain the server alive but without participating or engaging in any new actions related to OntoPortal. Concerns were shared in the Alliance about the sustainability and maintenance of this instance.

Newcomers and interested parties

Baptiste Cecconi (OntoPortal-Astro)

The [OPAL project](#) (tender from EOSC project OSCARS) aims to establish OntoPortal-Astro, a dedicated Semantic Artefact Catalogue for astronomy, leveraging the OntoPortal technology. This portal will integrate vocabularies and other semantic artefacts from various domains within the [ESCAPE cluster](#) of research infrastructures –including observational astronomy, planetary sciences, heliophysics, and particle physics– to enable semantic interoperability, FAIRness assessment, and enhanced data discoverability. OntoPortal-Astro builds on the expertise of the IVOA Semantics Working Group, expanding its vocabulary framework to encompass a broader scope of astronomy and related disciplines. This initiative is crucial for fostering interoperability and supporting the diverse needs of the astronomy research community within the ESCAPE ecosystem.

Beatrice Markhoff (CHPortal)

Cultural heritage community via the [ECHOES EU project](#). ECHOES started in June 2024 and will last 5 years (until 2029), with 51 partners from 15 countries. It aims to set up an open science cloud dedicated to cultural heritage professionals and researchers. In order to help the communities sharing existing and future knowledge, it is planned to set up a CH semantic resources portal based on OntoPortal. Important semantic resources in this community include CIDOC-CRM or the ones reviewed by the ATRIUM EU project. The goal is to bring together fragmented communities within the cultural heritage sector into a cohesive community, fostering collaboration and co-creation of new knowledge through the use of shared digital resources and tools, especially with respect to semantics. No technical installation or prototyping has been started so far. No convergence with the previous Clover installation of OntoPortal is to report, for now.

Yuriy Marykovskiy (TechnoPortal)

The TechnoPortal initiative is a semantic artefact catalogue for wind energy and technological sciences developed by Ost / Hevs in Switzerland.

Darren Bell (CESSDA)

Invited but not present. Express interest in adopting OntoPortal for vocabularies in the social sciences and humanities domain as currently served by the CESSDA infrastructure.

Clement Jonquet (LovPortal)

The project involves deploying a dedicated ontology portal for semantic web vocabularies using the OntoPortal technology, inspired by the [Linked Open Vocabularies \(LOV\) platform](#). This initiative, named **LovPortal**, aims to host, share, and serve vocabularies and ontologies critical to the Semantic Web community or that are currently stored on the LOV platform. In parallel, the project will contribute to enhancing OntoPortal by developing new features, such as automating usage report generation and improving platform configuration and administration. Additionally, the ontology upload process will be revisited to streamline its usability. The work will be conducted in collaboration with UPM (Universidad Politécnica de Madrid) currently running and administering the LOV platform. The LovPortal, by hosting and serving the W3C vocabularies or other standard metadata vocabularies can play an important role in the OntoPortal federation.

Action items

- ☐ Support the deployment of OntoPortal-Astro within the OPAL project.
- ☐ Support the deployment of CHPortal within the ECHOES project.
- ☐ Support the deployment of LovPortal within the FAIR-IMPACT project and in partnership with the LOV team.
- ☐ Support the deployment of SHS portal within the CESSDA infrastructure.

Summary of the topic sessions

Each section finishes with **Action items**. When the topic of the session was already addressed in the previous workshop 2022 and 2023 reports, we have reported section per section the progress with the following legend:

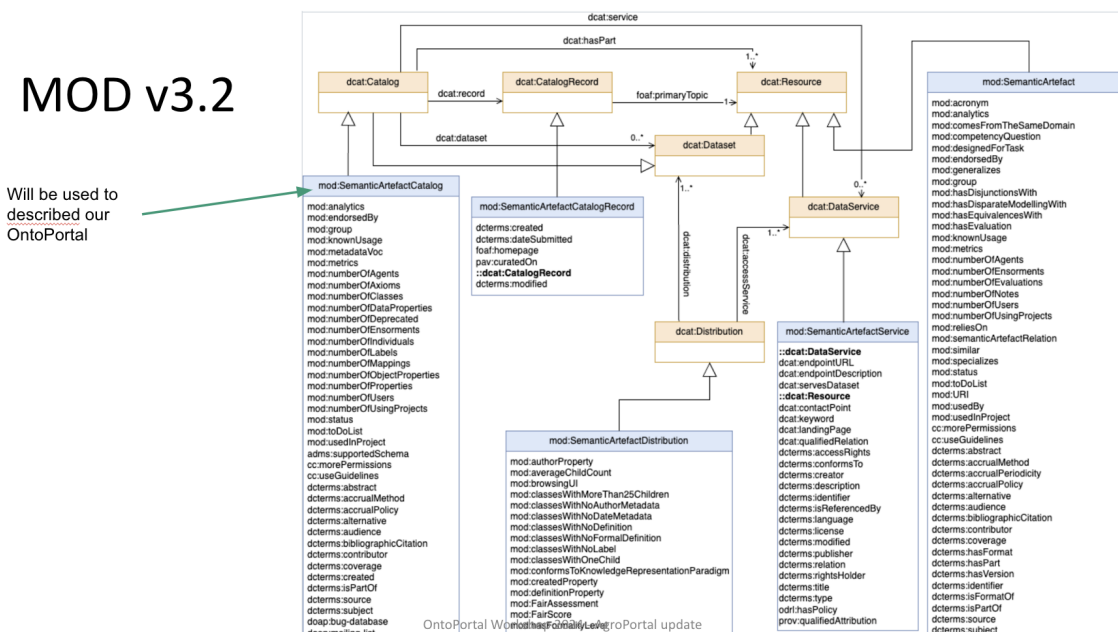
- Reported in 2022/2023 and achieved in 2023/2024 in either one specific portal or at the OntoPortal level.
- Reported in 2022/2023 not yet achieved but moved to GitHub for further discussion.
- Reported in 2022/2023 and not yet addressed or not yet ready to be moved to GitHub for discussion and realization.

OntoPortal related work within FAIR-IMPACT

Work done in the context of AgroPortal/EcoPortal/EarthPortal in FAIR-IMPACT:

- Results from a survey on semantic artefact catalogues (<https://doi.org/10.5281/zenodo.12799795>) featuring the most important technologies including OntoPortal was presented. In this review of 173 semantic artefact catalogues and 15 generic technologies, one may observe OntoPortal is the technology that is the second most used (after SKOSMOS) with 9 public repositories listed.

- The standardization effort for ontology metadata was also quickly reported, especially the release of MOD v3.2 (<https://github.com/FAIR-IMPACT/MOD>) which now offer a semantic object (mod:SemanticArtefactCatalogue) with identified metadata properties that can be used to semantically describe our repositories.



- The corresponding v1 of MOD-API (<https://github.com/FAIR-IMPACT/MOD-API>) was presented as a proposition from FAIR-IMPACT of a specification of a shared API that could be implemented by any semantic artefact catalogues to enable their interoperability and uniform access. The implementation of the MOD-API is the subject of an open call offered by FAIR-IMPACT in 2024: <https://fair-impact.eu/implementing-shared-api-for-semantic-catalogues> and it would be implemented within AgroPortal, EarthPortal and EcoPortal in 2025.

Action items

- ☐ Update the internal metadata model of several portals to be compliant with latest version of MOD, especially in terms of new properties provided.
- ☐ Implement the MOD-API in AgroPortal/EcoPortal/EarthPortal at least.

Connectors to OntoPortal

The goal of this session was to report on recent work done, especially in the context of FAIR-IMPACT T4.5, on developing connectors from research data repositories to OntoPortal. Typically these connectors allow users to semi-automatically pick-up controlled terms to annotate or index a dataset or record in a data repository. They often rely on the Search service but may sometimes also use the Annotator service. The session also established a comprehensive listing of all the connectors known

today, listed in the following tables. The 5-first were explicitly presented/demoed during the session.

(meta)data catalogue technology	Data catalogue concerned	Lead OntoPortal use case	Partner	Comments / observations
DSpace	https://data.lifewatchdev.eu/search	EcoPortal	Ilaria, Enrica, Martina, LW Italy	
GeoNetwork	https://metadatalifewatchdev.eu/geonetwork/srv/eng/catalog.search	EcoPortal	Ilaria, Enrica, Martina, LW Italy	Data Terra will also work with partners to do a connector to EarthPortal.
Dataverse	DataINRAE	AgroPortal	Dimitri, Sophie (INRAE-Dip SO) Clement	Generic enough to connect easily to another OntoPortal. Future plans will be to use the semantics of the vocabulary.
Not generic (deposit) Geonetwork (Harvest & catalog)	EasyData	EarthPortal	Christelle, Hélène (BRGM)	Deposit use case works. Next: harvest (directly with Geonetwork) and Search.
OpenSILEX	PHIS	AgroPortal	MISTEA lab + Llorenc	Works now in PHIS. Will be generalized to other instances of OpenSILEX. Will eventually be extended to take proposals back.
Dai:Si Portal (not generic): Python/Django + Wagtail	GFBio Data Catalog	BiodivPortal	Naouel, GFBio e.V.	Semantic search, Data Integration based on data standards and text/table annotation under development within NFDI4Biodiversity
CEDAR	https://datadryad.org/stash , https://osf.io/	BioPortal	BMIR	
REDCap		BioPortal	REDCap, Vanderbilt	Built into the popular REDCap (medical) data description system for

				nearly a decade, and is among the heaviest issuers of BioPortal API calls.
--	--	--	--	--

It was decided that this list should be moved to the documentation and/or website as this is a very compelling argument for external parties to adopt OntoPortal. The developers of these connectors could be invited to participate in a hackathon or technical event attached to the next OntoPortal Workshop.

Action items

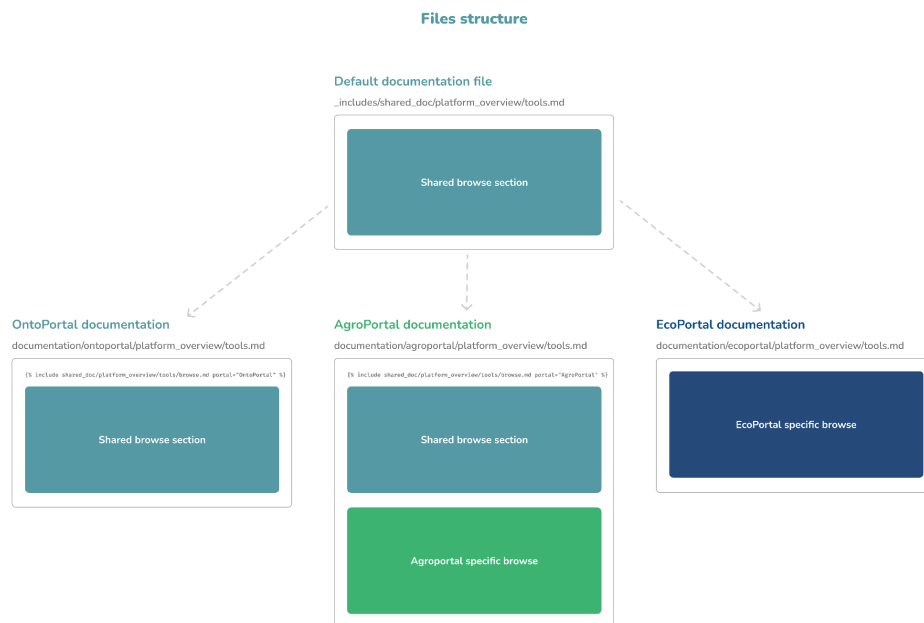
- ☐ Move the list of connectors to OntoPortal to the website and documentation.
- ☐ Promote these connectors into a shared publication.

Documentation and training materials

A dedicated working group has been actively reviewing and updating the documentation of OntoPortal and OntoPortal-based instances to produce a shared document (Users guide only) accessible at: <https://ontoportal.github.io/documentation>. The key idea is this documentation being modular by sharing common elements and can be customized by adding portal-specific content too.

The work done by the working group was to reuse what was existing as a foundation for the shared documentation. The EcoPortal team already had comprehensive user documentation which has been used and by pruning parts that were specific to EcoPortal, we then obtained the first shared table of content that would be used for all portals.

Each portal can now create its own documentation by weaving generic and portal-specific content, since it works like building blocks:



As long as portals share the same table of content titles, it is possible to switch between the documentation of different portals thanks to an array of switches at the bottom of the page. While any portal could change the table of contents on its own, any such changes to the core structure must be agreed by members of the Alliance before they will become visible in the shared documentation.

This page is available for these portals:



The table of content will then automatically adjust to match the current portal selected. Each of them can also add specific sections to the table of content which will not appear for other portals.

The documentation also includes a page on [how to configure and then contribute to the documentation](#). The overall approach to the documentation is similar to the one of the code repositories and the website: contributions are made as for any projects on GitHub. The current documentation includes mostly the new documentation that was made by and from EcoPortal. The generic part of this documentation is available under the “OntoPortal” tab.

A suggestion has been raised to include inline links from the portals to the corresponding documentation page but a generic solution has to be designed. Future plans include migrating [documentation from IndustryPortal](#), [AgroPortal](#), and [BioPortal](#).

Additionally, there are plans to transition API documentation to Swagger for improved usability.

Action items

- ☐ Refresh EcoPortal documentation after joint release and new UIs.
- ☐ Complete EarthPortal documentation.
- ☐ Move AgroPortal's specific documentation to new documentation.
- ☐ Move BioPortal's specific documentation to new documentation.
- ☐ Move IndustryPortal's specific documentation to new documentation.
- ☐ Publish these workshop reports on the website with a page/post describing the event.

Action from previous reports

- Video tutorial for the installation of OntoPortal.
- Tutorial on how to setup a development environment for a new OntoPortal developer
- Consolidate all the documentation available (deployer guide, developer guide, user guide) and set up a system that allows factorizing documentation between OntoPortal installations.

Mappings & SSSOM compliance

It is foreseen that OntoPortal's mapping repository¹ will need a major refactoring and evolution to answer present and upcoming community needs with respect to mappings. FAIR-IMPACT (and now a new [RDA FAIR mappings working group](#)) are offering a forum to discuss needs and trends related to mappings. During this session, we explored the current state and future directions for mappings in OntoPortal. The discussion was structured around two main questions: (1) What is the status of mappings in our systems? We reviewed what has already been implemented, what is currently in progress, and the opportunities presented by SSSOM and broader trends in mapping standards. (2) Who wants what for the enhanced mapping repository? This focused on identifying interested stakeholders, desired capabilities, and potential integrations with third-party mapping tools and repositories such as Cocoda, OxO, and MSCR.

Future plans include adopting a philosophy similar to our approach for ontologies—focusing on hosting and serving mappings rather than editing them. Enhancements would involve: enhancing visualizations (introducing new views or highlighting mapping availability), filtering mappings using metadata or scoring, enabling bulk downloads, and ensuring SSSOM compliance for both import and storage. These steps aim to position the platform as a robust and interoperable

¹ The "OntoPortal mapping repository" is defined as the collection of OntoPortal features that support mapping of semantic concepts, whether those concepts are within a given repository or external to it.

mapping repository supporting the community's evolving needs. The Alliance did not present a clear roadmap on this subject yet but agreed on considering this in the next cycle of funding (as such a work would require specific funding and use cases). The commitment to SSSOM is also clear to the group as SSSOM provides a rich, standardized model for managing mapping metadata.

Action items

- ☐ Plan the refactoring of the mapping repository and integrate this in the next funding round.
- ☐ Formalize issues and discussions on GitHub for a better follow-up of this topic.

Action from previous reports

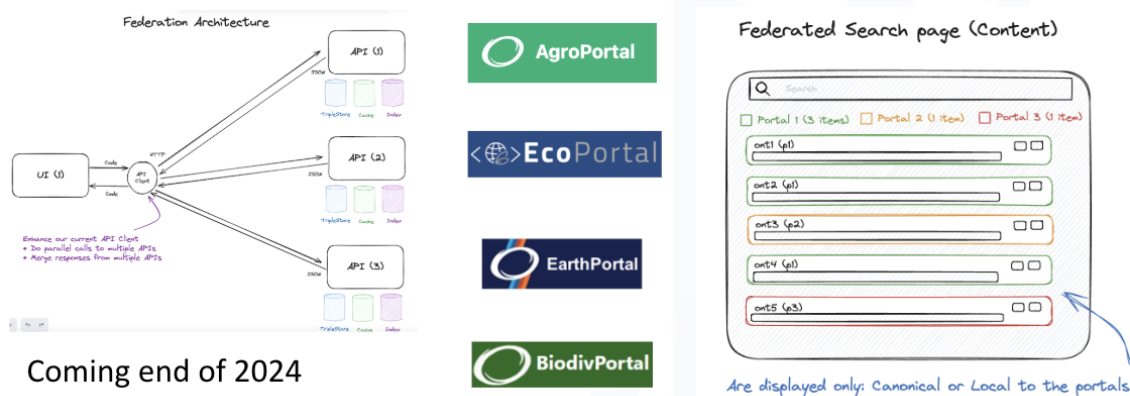
- Full support of SSSOM for the mapping repository.
- Implement better views for mappings including an evaluation view and an overview of mappings between any ontologies.
- Support bulk upload and download of mappings.
- Connect to third-party tools for ontology alignment. The appropriate way needs to be clarified as many tools are not generic or pluggable. Not necessarily process the cartesian product of mappings all the ontologies with all the other ones.
- Enhance the use of mappings inside the portal itself.
- Support evolution of mappings in the repository when mapped semantic resources evolve
- Make existing mappings selectable/viewable by their origin or any other mapping metadata (e.g., return only REST mappings, or only LOOM mappings).

OntoPortal federation

This session focused on progress and planning related to OntoPortal federation. Different approaches, currently being developed, to make semantic artefact catalogues interoperable were briefly discussed: (i) OntoPortal federation at the web application (user interface) level only (described hereafter); (ii) FAIR-IMPACT's MOD-API implementation and (iii) TS4NFDI's implementation of an API-Gateway. The Alliance presented the results of a first round of implementation addressing key functionalities such as federated documentation (cf. dedicated session), federated/harmonized user interfaces, portal descriptions, federated categories, federated browsing, and search. AgroPortal, EcoPortal, EarthPortal and BiodivPortal are involved in this federation and are planning a joint federated release toward the end of 2024 where each portal will consume content from each other. Technical discussions touched on benchmarks for search and browse, API key usage, caching, and downtime detection for portals. For federated categories, the choice of using the UNESCO Thesaurus to unify the categories across federated portals was presented.²

² Recent AgroPortal updates (v2.10) introduced a category/group view, while earlier versions enabled editing only.

Early work on federation of 4 catalogues (OntoPortal)



Future tasks include federating ontology metadata and different scenarios were presented including the one chosen (metadata pulled from an ontology-specific canonical portal) to best handle duplicates ontologies. Other challenges, especially regarding integrating BioPortal in this federation were discussed. Although BioPortal expressed that it was ok to get their content consumed by other portals, the group agreed not to consume content from BioPortal (even if technically possible) until BioPortal can itself be part of the federation and consume content from the other portals too. Overall, positive feedback was noted and the session maintained strong engagement and advanced critical elements of OntoPortal federation.

Action items

- ☐ Joint federated release of AgroPortal, EcoPortal, EarthPortal and BiodivPortal.
- ☐ Embark on including BioPortal, AstroPortal and LovPortal within the federation.
- ☐ Discuss and eventually plan OntoPortal federation (2nd round) including federated projects, agents, annotator, recommender, users, mappings, browsing, and loading.

Action from previous reports

- Implement a view to list (with linked logos) the public OntoPortal installation in the OntoPortal Alliance. A commitment to the Alliance means a presence on the page on all the portals.
- Better coordination when a semantic resource is hosted in multiple repositories in the alliance.
 - Avoid repeating the authoring of metadata from one portal to another.
 - Should be easy for final users to find content on different OntoPortal installations.
- Implement a federated search of content. Grab either statistics (number of hits) or even results (to produce snippets) in the querying interface.
- Implement a cross-portal metadata search service that will facilitate the identification of the semantic resources.

- Federated Recommender. When a query is passed to the recommender, select the option to query also the other Recommenders.
- Select a shared classification to define and harmonize our categories across portals.
- Align AgroPortal/EcoPortal/EarthPortal/BiodivPortal code level.
- Merge the DOI feature of EcoPortal for other portals with a rich metadata model.
- Implement a federated metadata search service in each portal and cross portal.

Connecting to 3rd party generative AI tools

This session focused on how OntoPortal and related ontology platforms can integrate with third-party tools, particularly Generative AI (GenAI) solutions, to enhance ontology related workflows. It began with an overview of key GenAI tools, including OntoGPT for extracting structured information from text, Dify for workflows such as retrieval-augmented generation and chatbot capabilities, and additional tools like Autology, NEON-GPT, LLMs4OL, LLMMap-Prelim, and DeepOnto, which support ontology enrichment, alignment, and automated learning. A demonstration followed, showcasing: (i) A Custom GPT prototype built with MatPortal's API and (ii) MatPortal's integration with Dify which leveraged AI for translating ontology labels, validating ontologies, aligning multiple ontologies, and developing chatbots capable of answering questions and performing ontology lookups using the MatPortal API. The session concluded with a discussion on advanced integration strategies, including training custom GPT-based assistants tailored to OntoPortal platforms and leveraging tools like Dify for creating AI-enhanced interfaces. Challenges such as curating training datasets and optimizing ontology representation for AI were also addressed, highlighting opportunities for deeper integration and expanded functionalities in ontology workflows.

Action items

- ☐ Follow up on the possibility to add a plug-in LLMs framework to OntoPortal.
- ☐ Experiment with using CustomGPT in other OntoPortal instances and possibly pay for it.

Ontology development lifecycle: OntoPortal for Protégé and VocBench

Status of VocBench and OntoPortal connections

The VocBench Suite is a powerful, web-based tool designed for the collaborative management of multilingual thesauri, ontologies, and code lists, that supports several features such as ontology and thesaurus management, multilingual support and validation workflow with a user-friendly interface. VocBench was integrated within EcoPortal back in 2020 to answer to the community needs of having a collaborative platform for creating and editing semantic artefacts. From then, EcoPortal has been part of the wide VocBench users' community and contributes to its technological enhancement.

VocBench comes with a series of functionalities, and several customisation have been done to the meet the needs of the scientific community:

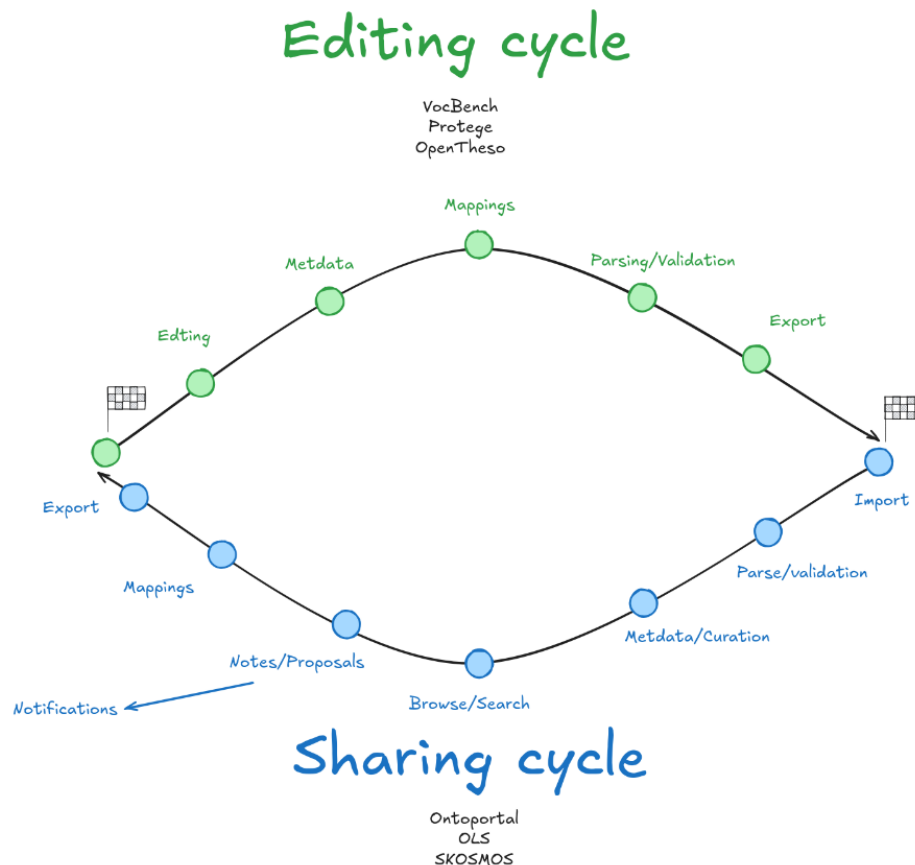
- **User management.** There is the possibility to create new roles and groups with different capabilities assigned from the administrators, providing a very high level of customisation. There are three types of users: admin, user and super-user, which is a specific development done for LifeWatch community and it is capable of creating projects, which is an admin-only feature.
- **Multilingual content support.** Each user is assigned with one or multiple languages in which he can make edits.
- **SPARQL editor.** Users can perform simple or complex SPARQL queries, with the possibility to download the results in several formats. The SPARQL queries can not be saved.
- **Versioning.** Users can create time-stamped data dumps of the semantic artefact, that can later be available through the same project. Each dump is stored in a separate repository, which is prevented from being re-written by the application.
- **History and validation.** The commit in the history can be inspected, showing the list of added/removed triples. Even if the changes that are made to the resource are visible, the option to revert to a previous version is not available. The validation workflow can be performed by users with specific capabilities, but it needs to be redesigned because it does not meet the community needs.
- **Alignment validation tool.** It's a third party tool MAPLE that is the orchestrator of the alignment task. The alignment is performed by the remote alignment platform. Users can save the alignments directly in the resource file or export it as RDF triples.
- **Semantic artefacts imports.** The import can be done also from catalogs like LOV, LOD, European Open Data Portal or OntoPortal.
- **Semantic artefacts export.** Specific customisations have been done to build a connector between VocBench and EcoPortal. The OntoPortal deployer checks that the data being submitted from VocBench satisfies the additional constraints of OntoPortal.

More details can be found in the public VocBench documentation (<https://vocbench.uniroma2.it/doc/user/>). Along with these functionalities some issues have been raised, such as the lack of a form for new project requests in the homepage, validator workflow to be redesigned, improve the UI to have a better user experience, activation of the SSO from users coming from EcoPortal.

Session reporting

Insights from the Rome meeting (May 15, 2024) on OntoPortal and VocBench collaboration were reviewed, emphasizing usability and shared functionalities. The session welcomed Matthew Horridge and Christian Kindermann from the Protégé team at Stanford. The discussion focused on the integration and development of VocBench

(described above) and Protégé within an ontology lifecycle that include OntoPortal. Key discussions addressed enhancing workflows between VocBench and OntoPortal, such as importing ontologies, aligning mappings, automating term proposals and translations, and deploying metadata profiles (MOD). The key idea was to discuss how the Alliance can respond to use cases interested in services and tools to manage the full life cycle of their semantic artefacts. Future implementations emphasized improving interconnectedness between these platforms, enabling streamlined ontology life cycle management through collaborative use cases.



Historical integrations, like the [BioPortal Import Plugin](#), were revisited, and the potential for re-establishing similar functionalities was explored, particularly for WebProtege, given the discontinuation of Stanford's support for Protege desktop. The session concluded by identifying opportunities to create a cohesive ecosystem, leveraging contributions from the Protege team and expanding collaborations to enhance ontology management and user experiences across semantic artefact repositories. No clear roadmap has been identified and the engagement of both VocBench and Protégé is at this time very limited.

Action items

- ☐ Continue exchanges about integration with VocBench.

- ☐ Continue exchanges about integration with Protégé.
- ☐ Formalize issues and discussions on GitHub for a better follow-up of this topic.

Action from previous reports

- Re-develop an OntoPortal plug-in for Protégé.
- Develop a basic link to an ontology editor specific to a semantic resource similar to what EcoPortal has done for VocBench (can be based on metadata attribute).

Community: notes, change requests, social

The session covered recent updates on BioPortal change request functionalities and integration with GitHub. The change request capability now supports renaming and obsoleting classes, and adding/removing synonyms. The entire change request suite of features continues to be enabled on a per-ontology basis.

The discussion then extended to a broader refactoring and unification of all the “social features” in OntoPortal such as Reviews, Notes (Proposal & Comments), Change Requests, Provisional Classes/Relations Notifications/Subscriptions, Reviews and Ontolobridge. Although Ontolobridge is now entirely gone, its corresponding need or *raison d’être* has not diminished. Notes were clarified as two distinct types: comments and proposals, prompting discussions on whether provisional classes/relations and note propositions could be merged. Similarly, the potential to merge change requests as a new type of Note was explored, alongside ideas for connecting with GitHub to display related issues directly. Reviews, although removed from the user interface, still exist in the backend, raising the question of whether they should be reintegrated. The session also examined the types of notifications required to support these features effectively. A possible integrated vision (mockup bellow) was presented and discussed.

AGROVOC (AROVOCRONIM)







 Watch (12)

Last update on February 2, 2023

Summary
 Concepts
 Properties
 Community
 Mappings
 Schemes
 Collections
 Widgets

Comment
 

Proposition
 

Subject

Leave a comment...




 Audience: **WorldWide**

Clement Jonquet

 2 Months ago
 substance unit

Duplicate modelling with classes and instances
 We found out that picomole, nanomole and micromole are duplicated as instances of the class "Substance unit" and subclasses of this same class.

2 Responds
 
 Respond
 

Syphax-bouazzouni

 2 Months ago

Revise the metadata fields available on the new "New submission" page
 We need to revise the metadata fields available on this page. It would need to at least offer the properties listed in the "versioning paper".

[View on github](#)


Bilel KHAL

 2 Months ago
 Textual entity

Duplicate modelling with classes and instances
 We found out that picomole, nanomole and micromole are duplicated as instances of the class "Substance unit" and subclasses of this same class.

2 Responds
 
 Respond
 

Hiba Djebabria

 2 Months ago
 Symbol

Duplicate modelling with classes and instances
 We found out that picomole, nanomole and micromole are duplicated as instances of the class "Substance unit" and subclasses of this same class.

2 Responds
 
 Respond
 

Leave respond ...

Action items

- ☐ Work toward uniformisation and harmonization of all social means.
- ☐ Possibly enable Reviews but based on an admin parameter (per ontology) and with possibility to an admin to review the reviews.

Action from previous reports

- Refactor the notes to be more modern and connected to the semantic resource developer practices.
- Support a unique, solid, and relevant term proposal process (currently OntoPortal has 4: "Add proposal", "Provisional classes", OntoloBridge and GitHub change requests).
- Refactor notifications (changes, notes, processing status, errors, new version, etc.). e.g. so ontologies/users will know when an ontology changes.

Alliance governance and funding

Funding and Project Updates

The session highlighted the funding landscape for represented OntoPortal projects. **BiodivPortal** is likely to receive renewed support from NFDI4Biodiversity (five years starting October 2025) and TS4NFDI (two years starting January 2025, with potential for extension). **EcoPortal** benefits from ITINERIS (EU PNRR funding until November

2025), FAIR-IMPACT (ending May 2025), and LifeWatch ERIC subcontracts. **AgroPortal** is transitioning to formal adoption by INRAE as a service, with ongoing support from FAIR-IMPACT (ending May 2025) and new projects like OPAL and FIDELIS starting January 2025. **BioPortal** requires a new funding proposal by spring 2025 as U24 (ending August 2026) and RADx (ending December 2025) approach their conclusion. **EarthPortal** anticipates support from EU project LUMEN, French RI Data Terra, CNRS INSU, and potential French ANR initiatives. However, the future of portals like MatPortal and IndustryPortal remains uncertain.

Alliance structure and governance

Discussions underscored the need to address licensing challenges and possible help to Stanford to sustain these efforts. Until now Stanford only has control on the licensing server and related information. The group discussed formalizing the OntoPortal Alliance as a legal entity, potentially through a non-profit organization. Models such as US's 501c3, a French association (1901 law) or an international AISBL (as in Belgium) were considered. It would enable hiring personnel, managing memberships, and accessing grants. A draft structure for a French association was suggested, requiring minimal overhead and aligning with ongoing funding initiatives: we emphasized the importance of creating the association soon to integrate it into upcoming ANR project proposals. Membership fees (maybe between 2000€ and 5000€ per year) were proposed as a sustainable mechanism for operational support, hiring a manager and/or lead developer with a focus on fostering collaboration among Alliance members. A formal membership status or procedure will then be necessary too along with a more formal decision process.

Technical challenges and alignment

The governance session also highlighted technical divergence among OntoPortal forks (e.g., BioPortal, AgroPortal, OntoPortal) due to a lack of consistent maintenance of the shared codebase. This has resulted in challenges with merging updates and avoiding redundancy in development. The group agreed on the need for better coordination to realign features, standardize processes, and improve the workflow for updates and releases. Use cases were identified to cater to diverse needs: AgroPortal's ability to evolve with new use cases versus Stanford's focus on a stable appliance for traditional users. The group also discussed establishing a shared list of ontology candidates. Something where each team could record semantic artefact of interest discovered. AgroPortal, EcoPortal, EarthPortal are already maintaining such a list of their candidates. LovPortal is going to have one and maybe too OntoPortal-Astro.

Upcoming opportunities

Several funding calls were identified as potential opportunities. These include the ANR (focused on FAIR ontologies lifecycle), FAIR-IMPACT's MOD-API implementation, Horizon Europe's research infrastructure and EOSC-related calls, and EU-USA

collaboration possibilities. These calls emphasize aligning with ongoing projects and fostering partnerships with key infrastructures like CESSDA, DataTerra, and LifeWatch ERIC. Participants stressed the importance of securing leadership and coordinating proposals to maximize the Alliance's impact and integration into broader initiatives. The group agreed on the fact that searching for funding of the whole Alliance is very hard, and that a better approach is to follow scientific/research motivations and use cases and embark as alliance members as possible.

Action items

- ☐ Establish a shared list of ontology candidates.
- ☐ Engage into the process of creating an (international) association for OntoPortal Alliance, under French status.
- ☐ Plan a meeting and discussion with Franz related to the use of AllegroGraph within OntoPortal instances. Plan for some communication on the subject.
- ☐ Address the big differences between the codebases so that contributions can be made by all to the master OntoPortal codebase on GitHub.
- ☐ Do more communication and outreach: presentations, tutorials, workshops etc.
- ☐ Organize the next OntoPortal workshop in Berlin in Fall 2025.
- ☐ Decide if another focus meeting (like the Federation meeting of April 2024) is required and, if possible, consider the participation of the whole core group (including BioPortal).
- ☐ Plan for a 2 days hackathon before the next OntoPortal Workshop for developers doing database connectors. Maybe create a task force on this.
- ☐ Establish a mechanism for tracking funding opportunities and appropriately coordinating responses.

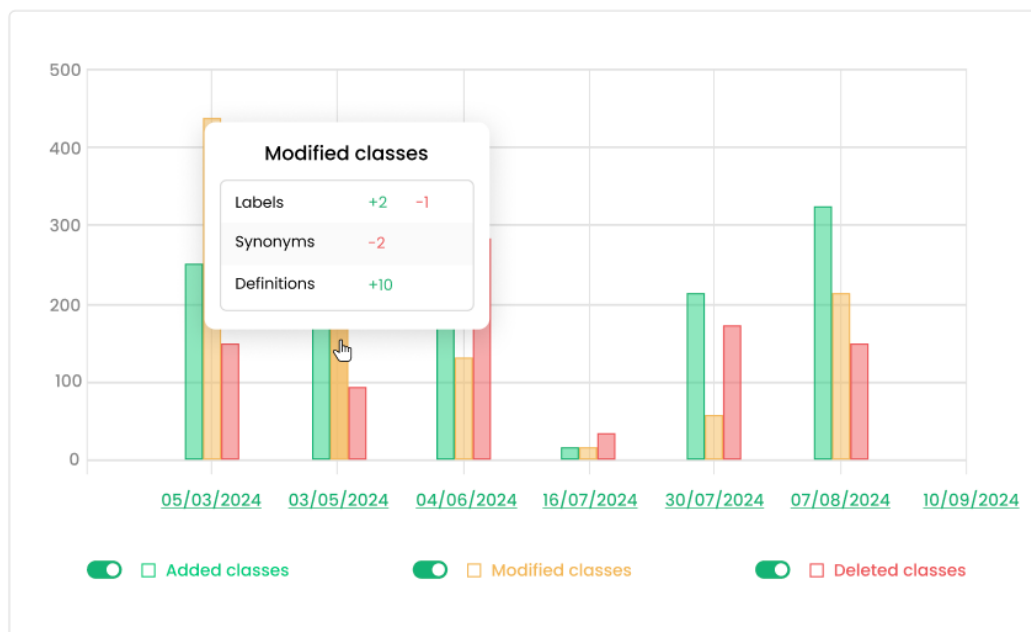
Ontology versioning/diffs

The session focused on approaches for managing and representing ontology changes, comparing existing implementations and tools. The discussion addressed models like KGCL (suited for OBO and partial OWL support), ROBOT, the Change Ontology from ContoDiff, and the historical ChAO ontology from Stanford-BMIR. Current limitations include the lack of direct integration between “diff” data and submission objects in OntoPortal (e.g., broken links to diff files). Tools like Bubastis were identified as needing enhancements, such as native SKOS support, to align better with OntoPortal's requirements.

	OAK	contonto_diff	Bubastis	Robot (owl-diff)	
Performance (EDAM ontology, 3.2M)	⚠️	✅	⚠️	✅	
Change types coverage	✅	✅	❌	⚠️	
RDF	✅	❌	❌	⚠️	
Maintained	✅	❌	❌	⚠️	
Improvable	⚠️	❌	✅	⚠️	
Support KCGL model	✅	❌	❌	❌	
Support SKOS	⚠️	❌	❌	⚠️	
Used parser	RDFLib(Python)	OWLAPI(Java)	OWLAPI(Java)	OWLAPI(Java)	

Proposed iterations for improvement include: (i) enhancing the existing diff mechanism with a user-friendly user interfaces (started) and (ii) evaluating alternative tools, storing change data in triple stores, and expanding RDF-based representations of diffs to include detailed statements about removed or modified URIs. These plans aim to standardize change history tracking and facilitate its integration into OntoPortal's broader ecosystem.

Diff



Action items

- ☐ Implement backend service and user interfaces for the current diff files generated by Bubastis

- ☐ Decide to adopt and enhance Bubastis or switch to another tool.

Action from previous reports

- Better support ontology diffs and improve the perception of what items have changed between versions and overall ontology evolution.

Reimagine the Projects feature

A discussion to upgrade the Project feature has been started in [this issue](#) to improve their metadata and linked capabilities with other components of the portal, but possibly to some external resources as well.

Currently, the page dedicated to Projects on the web interface needs a visual overhaul to fit the latest website design. Then, while OntoPortal should not become a project database, having additional metadata about it (logo, contacts, ...) would enhance the ability to link these with the corresponding ontologies, effectively improving the overall visibility of the resources produced or used during projects and so their compliance to FAIR principles. That provides more ways to filter ontologies within the portal to easily find which ones are connected to a project.

There are plans to also create connectors with external project databases (such as Cordis for european projects or ANR for french projects) to automatically populate the metadata and ease the recording process. Another feature outlined in this discussion is to add the possibility of providing metrics for external datasets using ontologies hosted in the portal.

This work is the subject of an internship expected to begin in early 2025 hosted by the EarthPortal team with assistance from the AgroPortal team.

Action items

- ☐ Redesign the Project web page to fit the current design
- ☐ Add new metadata fields to the current Project model
- ☐ Add a new filter to search ontologies by project
- ☐ Create connectors to external project databases to automatically populate metadata

Action from previous reports

- Link OntoPortal semantic resources to external datasets using or annotated by the semantic resources.

Technical meeting

CI/CD and Dockerization

While Dockerization offers benefits for development environments, its readiness for appliance production remains uncertain. Suggestions included testing production BioPortal in containers at Stanford to gather insights, with a potential focus on running backend services on VMs and UI/API applications as containers. There was consensus

to release version 4.0 of the appliance promptly, balancing immediate needs with future enhancements. Participants agreed on the importance of frequent appliance releases (at least once a year) to ensure alignment with the evolving codebase. The ncbo master should be used to compile the OntoPortal virtual appliance at least once a year (ideally more frequently). The *ontoportal-lirimm* master will be used only if requested by newcomers in the Alliance.

OntoPortal upstream state and code alignment

There is a clear need for convergence in the codebase and consistency in contributing updates to the upstream. Plus, appliance version 4 is a significant goal, incorporating new configurations such as Ubuntu, AllegroGraph by default, and enhanced deployment tools like Puppet and Packer. Migration documentation, updates for Ruby 3, and a choice of triple stores were identified as critical needs for smooth transitions. The alignment of the UI=frontend (2 GitHub repos) and API=backend (5 GitHub repos) components across portals revealed substantial divergence, particularly in the UI. Collaboration on pull requests, prioritization of shared efforts like UI testing, and careful integration of reusable components were proposed to reduce fragmentation. The API was found to be reasonably aligned as a collaborative effort from AgroPortal/BioPortal was done when integrating subsequent changes related to supporting multilingualism including an important refactoring of Goo for better SPARQL queries support and better code maintenance. This makes the group think we can target a unique OntoPortal backend software in the mid term. It was agreed that we could take different approaches for the frontend and the backend to realign the ncbo and *ontoportal-lirimm* backends, both based on the OntoPortal upstream. The frontend is more complex and will require a different approach, based on demanded pull requests with the assurance to be merged

Starting in November, the AgroPortal team will initiate self-created pull requests (PRs) to the OntoPortal backend repositories. The goal is to closely align the OntoPortal upstream with the *ontoportal-lirimm* master branch. Key changes, prioritized by the BioPortal team, include (not ordered):

- Support for multiple triple stores
- Ontology metadata and agents
- SKOS support
- Enhanced search indexing and integration with a new Solr version
- Mappings improvements
- URI management
- Single sign-on (SSO)
- Advanced analytics
- Ontology instances

The BioPortal team will work to synchronize their *ncbo* repositories with the OntoPortal upstream. The AgroPortal team will contribute PRs to the OntoPortal frontend repositories based on RADx requirements. These focus on:

1. Maintaining a shared architecture (e.g., transitioning to *esbuild*).
2. Adding view components (e.g., *gem* and configuration features).

The BioPortal team will merge these frontend PRs into their *ncbo* repositories and regularly submit PRs to the OntoPortal frontend upstream to ensure it stays aligned with the *ncbo* master branch.

The upstream codebase (both backend and frontend) may not be fully tested as a unit, or represent the exact configuration used in any production environment. Consequently, potential OntoPortal technology users may be directed to the BioPortal team for downloading and deploying a stable virtual appliance; or to the AgroPortal team for deploying a community portal using their latest code. Each team is responsible for ensuring their respective frontend and backend repositories (*ncbo* and *ontoportal-lirimm*) work together as needed in their production environment.

The session concluded with a shared commitment to resolving codebase divergence, enhancing technical alignment, and ensuring consistent contributions to the upstream. These efforts aim to strengthen the OntoPortal ecosystem and its collaborative foundation.

Action items

- ☐ Quickly (october) resync the OntoPortal upstream (both backend and frontend) with the *ncbo* repositories.
- ☐ Include AllegroGraph as the default triple store, with 4store as an optional fallback.
- ☐ Transition to Ubuntu for appliance development.
- ☐ Release appliance v4 with the *ncbo* codebase.
- ☐ Re-establish frequent appliance releases (at least once a year) aligned with the *ncbo* codebase.
- ☐ Initiate the next use cases using appliance version 4.
- ☐ Align API code (5 repos) as much as we can to converge toward a unique OntoPortal backend software adopted by all current active public portals.
- ☐ Align UI code (2 repos) as much as we can at the architecture and technical choices level. Leave to each portal the final choices of look-and-feel and activated UI components.
- ☐ Enforcing and adding automated testing for the generated distribution.

Action from previous reports

General

- Refactor the ontology submission process.
- Multilingual support.
- Better support of SKOS (something more aligned to what SKOSMOS does).
- Make the UI easily and more customizable.
- Creation of different user roles (curators, admin of a group of ontologies).
- SHACL shapes-based validation of semantic resources. Semantic validation functionality that will automatically check a semantic resource with respect to identified guidelines.
- Implement a read-only version of OntoPortal i.e., an installation with no user contributions (ontologies, projects, mappings, notes, login).
- Better support of ontology imports and re-uses of objects (classes, properties, instances, concepts) from other semantic resources.
- Better handling of ontologies that are knowledge bases (i.e., OWL model with many instance data)
- Better support of the Annotator as many many data are still in text form.
- The portal needs to facilitate ontology competency questions e.g., by showing the type of questions/queries an ontology allows to answer.

Metadata

- Refactor the Summary page for portals with a rich metadata model.
- Whatever the metadata model, consolidate the automatic metadata extraction and generation to ease the management/curation of metadata.
- Better capture semantic resource status (manage obsolete resources).
- Refresh cache after ontology submission for every other user to see a new submission.
- Enable a parameterizable metadata model.
- Implement an incrementally augmentable minimum metadata model allowing any portal to incorporate new metadata properties 'one-by-one'.
- Whatever the metadata model, find a way to import metadata from external sources e.g., a VOA file or OBO Foundry YAML description.
- Better description of the OntoPortal pages (not only Summary) with Schema.org for better indexing by Google.
- Evaluate the best way to adopt a more advanced metadata model in BioPortal (ideally compatible with the rich model of AgroPortal, and with requirements for FAIR ontologies).
- History view for the evolution of a semantic resource (historical influences between ontologies view as a graph)
- Connect to tools to Vidoco to produce HTML doc for (small) semantic resources

DevOps and environment

- Docker container-based setup/installation of OntoPortal for developers.
- Implement and describe an easily deployable environment test.
- Regression testing automatically on GitHub code updates.
- Implement data (ontology) migration scripts.
- Better support versioning of code and identification for which code is currently running on which appliance.
- Move to a more generic (and GDPR compliant) tool for analytics beyond Google Analytics.

- Support multiple open-source triplestore backends.
- Clean up a lot of code bits that are workarounds for 4store performance issues.
- Implement full SPARQL 1.1 compliance to support any triplestore (if not complete already).
- Implement a triple data migration script to move data from a 4store to another triplestore.
- Add monitoring tools to OntoPortals.
- Create a SPARQL query editor/viewer. Minimally plug in the viewer of the underlying used triple store (4store and AllegroGraph have one).
- Address the end of support of CentOS.
- Create an option to enable the use of a separate public SPARQL triplestore.
- Design other views to interact with the SPARQL triplestore e.g. RelFinder, AdvancedSearch, etc.
- Implement an entry point to a unique knowledge graph with the content (not duplicated) of all OntoPortal installations.
- Ultimately OntoPortal could be installed by “pushing one button” within the appropriate research infrastructure

Other technical aspects

- Update Bootstrap version.
- Upgrading to Rails 7.
- Migrate and define all our reusable UI components as view components and use a lookbook framework for previewing/testing the UI components.
- Optimizing process for large ontologies.
- Eliminate the old JS libraries.
 - Adopt Stimulus and Hotwire to avoid JavaScript.
 - Adopt ViewComponents as reusable view components.
 - Update our JS so we aren't forced to use jquery.
- Low-level caching in Rails may be the solution for a better caching in the UI (standard Rails cache will not work for our case).
- Refresh the widgets.
- Refactor multiple AJAX calls.
- Break up the ontology parsing process into smaller steps, more modular and traceable and executable in parallel.
- Use a better queue management tool like Sidekiq.
- Avoid looping over classes multiple times.
- Revise the mechanisms to deal with obsolete classes (branch and custom property) and see what's going on with the obsolete classes graphs.
- Move the generation of mgrep dictionary process to once per day.
- Audit on graphs cleaning in the triplestore when a submission is deleted.

Conclusion

The OntoPortal Alliance is at a critical juncture, requiring strategic decisions to strengthen its governance, align technical development, and secure sustainable funding. While challenges like codebase divergence and licensing persist, the 3rd OntoPortal workshop underscored the Alliance's potential to foster collaboration,

address community needs, and leverage upcoming opportunities to enhance its role in ontology and semantic artefact management. OntoPortal's strong position as a widely used ontology repository codebase, and the increasing number of deployed and planned public OntoPortal instances, reinforces both its actual importance and its potential to meet those goals.

The workshop concluded with a strong consensus on the next steps to advance both governance and technical alignment. Establishing an association was identified as a feasible first step toward formalizing the collaboration, enabling resource pooling, and ensuring operational sustainability. On the technical front, participants recognized the need to converge codebases and prioritize shared functionalities, such as API-level federation and connectors, while highlighting the importance of mutualizing not only development efforts but also community support for publishing and utilizing semantic artefacts. Finally, participants agreed on the necessity of securing targeted funding for high-priority features such as refactoring the Annotator and the Mapping repository.

The Alliance is concerned to ensure that all members benefit from the collaboration without adding undue burden. The workshop concluded with optimism about the alliance's capacity to build a more integrated and impactful ecosystem. The idea of a dedicated mid-term meeting focused on specific topics was well-received, emphasizing the importance of maintaining regular communication and shared progress. Plans for the next OntoPortal Workshop, potentially in Berlin at Fall 2025, and hosted by the BiodivPortal team, reflect the Alliance's commitment to continuity and inclusivity.