HAL open science

OntoPortal Workshop 2022 Report

Clement Jonquet, John Graybeal

▶ **To cite this version:**

Clement Jonquet, John Graybeal. OntoPortal Workshop 2022 Report. Université de Montpellier; Stanford University. 2022, 16p. hal-04087929

HAL Id: hal-04087929

https://hal.science/hal-04087929v1

Submitted on 11 May 2023

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.

Distributed under a Creative Commons Attribution 4.0 International License

OntoPortal Workshop 2022 Report

Montpellier, LIRMM, September 26-29th, 2022

Clement Jonquet, John Graybeal
and the members of the OntoPortal Alliance

Participants: Mark Musen, John Graybeal, Jennifer Vendetti, Michael Dorf, Tim Redmond, Alex Skrenchuk, Clement Jonquet, Syphax Bouazzouni, Nicola Fiore, Xeni Kechagioglou, Ilaria Rosati, Agis Papantoniou, Alexandru-Aurelian Todor, Hedi Karray, Arkopaul Sarkar, Abdeladoud Rasmi, Christelle Pierkot, Naouel Karam, Ralph Schäfermeier, Yann Le Franc

Summary of the event **2**

Discussions **2**

OntoPortal philosophy 2

Feature requests identified by the partners 3

Metadata 3

Mappings 4

Federated access 4

DevOps and environment 4

Feature requests identified during the public track 5

Decisions & next actions **5**

Short term action items 6
Topics and leaders identified for the next few months 6
Technical points and bugs 6
Ontology parsing 7
Triplesore evolution 7
Caching and UI evolution 7

Observations 8

Summary of the event

The OntoPortal Alliance is a consortium of research and infrastructure teams dedicated to promoting services for ontologies and other semantic resources—in science and other domains—based on the open, collaboratively developed OntoPortal open source software. Teams in the Alliance develop and maintain several openly accessible semantic resource repositories. This includes BioPortal, the primary and historical source of OntoPortal code, and also AgroPortal, EcoPortal, MedPortal and MatPortal. List of teams on GitHub

The OntoPortal Alliance members and other partners gathered from Sept. 26 to 29th, 2022 in Montpellier for the 1st OntoPortal Workshop. The meeting gathered around 20 persons including management, research and technical profiles. The 2022 workshop's main goals were to consolidate the OntoPortal Alliance organization and shared agenda, and to inform and engage the larger scientific community interested in the OntoPortal platform.

The program included several sessions (technical, content, management) as shown in the program overview, and a public session attended by 35 participants (including 10 persons from the workshop). The public session included a technical OntoPortal installation tutorial.

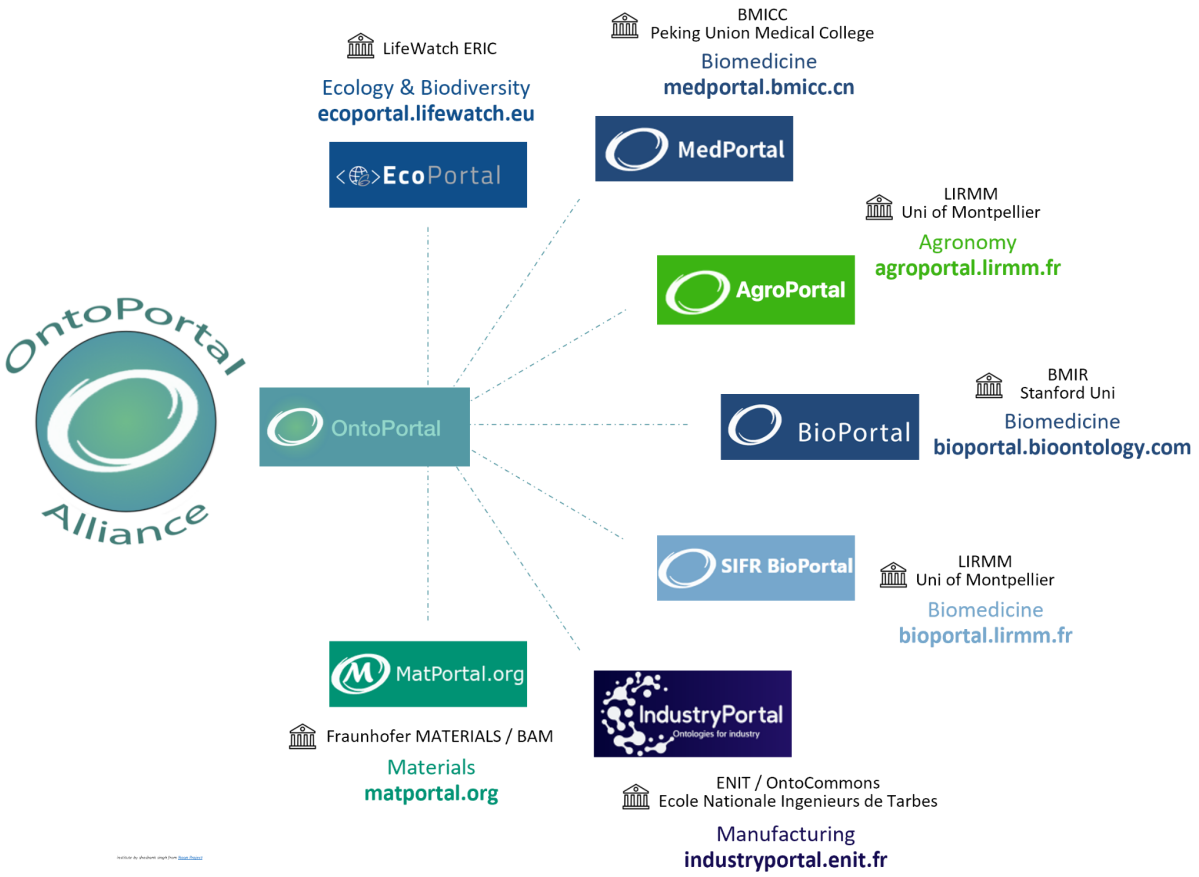
In preparation of the event, a survey gathered the partners' perceptions and needs with respect to their involvement in the group. Each team submitted a response for the portal (or portals) developed by their team. See survey responses. In addition, each partner's presentations are available here.

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.

Discussions

Output from the team presentations and survey responses



BioPortal

BioPortal is the foundational and most complex project of all the OntoPortal deployments. Started in 2005 in partnership with what became the OBO Foundry, the platform has over 1 million API accesses per day, a complex infrastructure and many users whose activities are hard to follow and track within the platform. The software has been developed over 3 major generations and the current one (based on 4store and Ruby/Rails) has been the baseline for the OntoPortal Appliance since 2012. However, mostly for scalability and robustness reasons, BioPortal's infrastructure is seen as very complex by most of the groups in the Alliance, as they do not (yet) have these requirements. In some situations, the software relies on third party software and while that has proved very useful, it creates dependencies that can be challenging to support.

Some features or aspects, while still of interest for other OntoPortal installations, have been experimented with and sometimes discontinued; the Alliance may learn from this experience. We list several instances: (i) visualization of ontologies is still a big need, although it has not changed in BioPortal in 10 years; (ii) notes and community feedback never really caught on although social and collaborative aspects of building ontologies has increased; (iii) the Resource Index was a good way to connect BioPortal to datasets/databases but the system became too complex to manage with ever changing data repositories whose APIs needed to be continuously updated. Other features, like proposals for changes to terms, have had several implementations and are still being discussed and developed today.

The software was and still is developed with a funding-driven and use-case-driven approach and features were developed to answer specific ontology project or community needs. Despite the enormous service to the community, BioPortal has not yet found a sustainable “infrastructure-level” support model while staying open and free for all scientists. This point is something to keep in mind also for other projects in the Alliance.

Today BioPortal’s team is very happy to see the re-uses of the technology originally developed for BioPortal in other OntoPortal installations; however the team and processes are not yet ready for the OntoPortal repository to be the main source of code and other resources to the Alliance. This would allow contributions of the OntoPortal code base to be a source of improvements for all the deployments, including BioPortal.

AgroPortal

AgroPortal was the first public reuse of the NCBO technology (before it was branded as OntoPortal) and demonstrated that it was possible (technically and scientifically) to have other instances of BioPortal that will reach out to other/new communities. AgroPortal presents itself as a generic tool for any kind of semantic resources (or artefacts) related to agri-food, but does not abandon the word “ontology” which directly connects to research activities. The project, started as a side project of the SIFR BioPortal (a dedicated biomedical instance for French use cases), has taken over but the team still keep the development approach that a feature has to be generic, domain-agnostic, and easily activated and parameterized in different OntoPortal installations.

The AgroPortal tool is being developed in the context of other research/scientific projects, but without (yet) any specific infrastructure or recurrent support. The team requires three roles: management, curation/outreach and development. This model might be applicable to other Alliance partners too. The project heavily curates the metadata of semantic resources in the system (metadata on which multiple new features like Landscape view and O’FAIRe rely), and tries to engage with the resource developers. AgroPortal has implemented many upgrades (mappings, annotation, and metadata being 3 examples), but the software (and API) is kept backward-compatible with Bioportal's main branch so that it can be merged back. The developers use pull requests to share back the code with BioPortal’s team.

MatPortal

MatPortal was created in 2021 to accelerate the development and reuse of ontologies in the Materials Science community. It is managed by a committee currently consisting of representatives from the Fraunhofer Group – Materials and BAM (The German Federal Institute for Materials Science and Testing) and supported by various German research projects from the Platform Material Digital and NFDI Matwerk. MatPortal is currently used as a repository for publishing materials science ontologies developed in research projects and from the IOF Materials Science and Engineering Workgroup, as well as offering crucial functionalities for ontology development such as visualization, search, mappings, and annotations. Further developments include specific features to support the development of standards specific ontologies in alignment with the Smart Standards initiatives.

EcoPortal

Created in 2019 and based on a study started in 2017, EcoPortal is supported by LifeWatch ERIC and updated through LifeWatch’s Italian national node, which gives the platform visibility and sustainability within the community of Ecology, Biodiversity and related domains. Beyond the main OntoPortal functionalities, the platform has developed a new function that permits creators/authors to request and obtain a DOI for hosted semantic resources. For this reason, EcoPortal has incorporated mandatory

metadata fields from the DataCite model (LifeWatch ERIC is a member of DataCite). EcoPortal offers also the ability to use an internal instance of the semantic editor VocBench, accessing it from the Tools and signing in with the same credentials. The offered feature involves direct deployment in EcoPortal of semantic artefact versions edited in VocBench, providing in this way a quite complete solution for their lifecycle.

Others

NFDI4biodiv

NDFI4biodiv is a node of the NFDI infrastructure projects in Germany. They are interested in adopting OntoPortal for developing the next generation of the GFBio terminology service. The convergence and partnership with LifeWatch EcoPortal are under discussion, especially on the editing aspects and common content coordination. The team has great experience (and ideas) with mappings of semantic resources and terminology evolution. A new developer will join starting January 2023, and the plan of adoption is validated by the NFDI4Biodiv consortium.

Cogni.zone

As an SME that has installed OntoPortal for client(s) e.g. Total Energies, Congi.zone has a different perspective on possible use cases and clients. The company can help reach out to different communities than what surrounds us (e.g. European Parliament). Technically speaking, Cogni.zone has installed the appliance using Microsoft Azure cloud; security was identified as an issue not easy to address. Around the company's ecosystem there are many opportunities around SKOS; the company works with Sparnatural to address some of these questions.

IndustryPortal

To alleviate the proliferation of semantic artifacts in the industry domain, a common ontology repository for industry was been established to serve the community with a long term permanent repository for FAIR semantic artefacts. Project Lead Hedi Karray's work around this started in 2006 with ontologies for industry and manufacturing, when he started the Industry Ontologies Foundry. Now, OntoCommons is an EU project to move semantics forward in the industrial domain, and OntoPortal is offering the possibility of semantic solutions. In this domain there is a strong motivation on making things interoperable so there is an important focus on aligning the semantic resources. For the moment, IndustryPortal runs the ontoportal-lirmm code branch with support of LIRMM in order to be able to do FAIRness assessment.

EOSC Pillar and FAIRCat

A specific instance of OntoPortal running in the context of the H2020 EOSC Pillar project was presented. This installation was a demonstrator. eScience Data Factory presented also the FAIRCat tool created to harvest the content of any semantic artefact catalogue (including OntoPortal installations but not only), describing their metadata with a DCAT-based extension metadata model and then publishing them back automatically in a web server following the FAIR Data Point specification.

EarthPortal

EarthPortal was not explicitly presented as Christelle could not join the meeting. It was mentioned as another coming (2023) installation in the earth sciences domain. The initiative is taken by the French Data Terra research infrastructure in the context of the FAIR-IMPACT project. One developer open position is available.

Survey outcomes

The first observation is that people responded thoroughly to the survey, which included 50 (!) questions, 19 of them free text responses. Most questions, including free-text items, received responses from all 9 respondents.

Most choice questions included at least one response that received over 50% of the answers. We can use these answers to construct **an image of a 'typical' OntoPortal Alliance deployment**:

> a public community repository where anyone can contribute ontologies, with more than 50 ontologies (eventually), less than 25% of them exclusive but more than 50% unavailable in any other OntoPortal repository; multilingual content (77%) and corresponding interest in supporting it; ontologies are mostly added by admins even though users can also add them; performing significant or moderate content curation; used by a few groups and users, maybe dozens to 100; heavily focused on web UI interface; having a 1-to-3-person team; emphasizing ontology curation, technical development and operations, and project management; deploying OntoPortal locally in a self-hosted server (single computer), with a recent version of the system; running default code or the forked code, but developing and deploying code additions locally; and in agreement that code should be publicly shared in Git in a common code base.

There was a strong commitment to the OntoPortal code base for future work, and most leaned toward the intent to find funding and organizational support for the Alliance.

In the survey, we requested and received detailed information about several facets of the deployments (these answers will be documented on the OntoPortal Alliance site). The most important reason people wanted to run an OntoPortal instance was the value of running a community-specific ontology repository (mean=4.1, smaller numbers meaning 'more important'), while the least important reason (mean=6.2) was BioPortal's reliability. Most have not determined any policy for adding ontologies to the collection, and had relatively little outreach (it is likely too soon for many deployments). There was a lot of interest in adding diverse features, and several responses alluded to improving the ability to re-use ontologies in various ways. Many groups expressed interest in leading a shared development action. The fact that the triple store might not remain free (for example, with Bioportal substituting Allegrograph for 4store) only bothered 2 responders.

In governance, many saw the key methods of working together including regular general meetings, topic-specific meetings, and open-source code sharing practices. But all the methods for working together received at least 2 votes.

Regarding concerns and issues, many answers referenced documentation; versioning; the technical elements Ruby/RubyOnRails and 4store; and deployment idiosyncrasies and difficulties.

Opinions on these questions were fairly evenly split: 'Should we have a shared approach for where ontologies should go?' and 'What are the most-used features of the OntoPortal installation?'

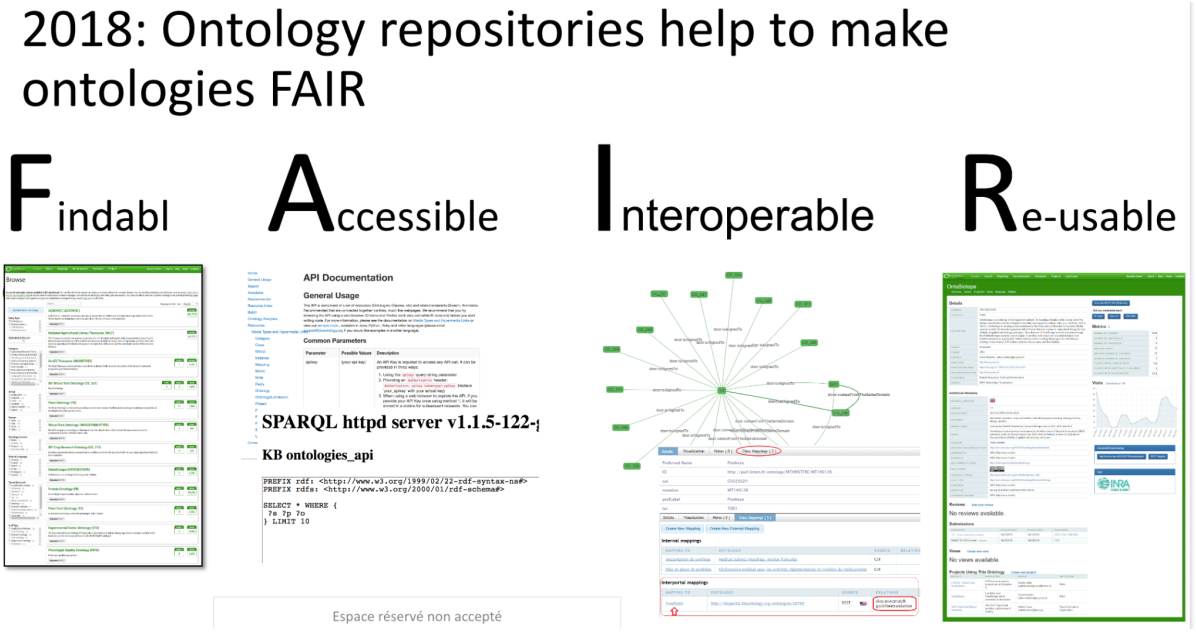
The complete survey results are available.
A synthesis is provided in the following table.

Aa Name	Bioportal(Xiaolin Yang)	IndustryPortal (Arkopa...	Bioportal (Timothy Red...	GFBio (Naouel Karam)	Agroportal (Clement jo...	EcoPortal (NICOLA FIO...	Cognizone (Agis Papan...	MatPortal (Alexendru)
Size	Ontologies > 50 with < 25% exclusive, dozen of users	Ontologies 10-50 with > 75% exclusive, dozen of users	Ontologies > 50 with < 25% exclusive, reference platform for a large community (more then hundred of users)	Ontologies 10-50 with < 25% exclusive, hundred of users	Ontologies > 50 with 50%-75% exclusive, hundred of users	Ontologies > 50 with 25%-50% exclusive, dozen of users	Ontologies > 50 with < 25% exclusive, only my own group project	Ontologies 10-50 with < 25% exclusive, dozen of users
Team	?	1-3	> 3	1-3	1-3	1-3	> 3	> 3
Interested in multilingual support	✅	⛔	✅	⛔	✅	✅	✅	✅
Coordinate to a common sharing ontology approach	✅	✅	?	⛔	✅	⛔	⛔	
Used features	Metadata description, Visualization / browsing, Ontology identification, Concept search and annotator	Metadata description	Annotator, Visualization / browsing, Concept search	Concept search, Mappings	Metadata description, Visualization / browsing, Ontology identification and selection, Concept search	Visualization / browsing, Ontology identification, Concept search and Metadata description	Concept search, Community related features, Metadata description, Mapping utilities	Concept search, Mapping utilities, Ontology identification and selection
Working on or plan to do	plan to use Ontoportal to help researchers to descript their experiments with ontology terms	A la carte ontology reuse (multiple ontologies with internal mapping) + OntoCommons ecosystem (Harmonized ontology stack from multiple to TLO to DLO)	fewer bugs, faster ingestion, new backend, support for git based ontology pulls, community support	A Docker container , Ontology evolution mechanisms and services, Ontology matching	A complete mapping framework. A complete user feedback and notification system (a la GitHub). Better graphical visualizations. Better interoperation with other semantic web tools and NLP tools (Annotator). A semantic search mechanism to related databases. Better SKOS support. SSSOM mapping compliance. FAIRness assessment. Better metadata model. Metadata curation procedures and interfaces.	Better support for SKOS, Integration with ORCID, Vocbenc, Datacite, Metadata set	Management of SKOS taxonomies, Azure compatible application image	Linked Data/RDF Instance Publishing, Mappings, Statistics, Data Pipeline, Dev/Sec/Ops
Centralized repository to all contribute together to the same codebase?	3/5	3/5	5/5	4/5	5/5	4/5	5/5	5/5

Aa Name	Bioportal(Xiaolin Yang)	IndustryPortal (Arkopa...	Bioportal (Timothy Red...	GFBio (Naouel Karam)	Agroportal (Clement jo...	EcoPortal (NICOLA FIO...	Cognizone (Agis Papan...	MatPortal (Alexendru)
Code changes	No, only a few cosmetic	yes	yes	No	yes	yes	yes	no
Obstacles encountered in deploying (and maintaining)	Hard to find approaches to the problem when we met difficulties.	Different versions and code base	It is a big system with faults and many dependent users.	Deploy our services on a Kubernetes Cluster	Detailed documented procedures. A mix expertise of devops and dev and semantics is needed and its rare. Instability is hard to track and monitoring is a bit chaotic. Hard to know the machine/server requirements in advance.	Be alligned with the Ontoportal version update	Azure compatibility (required specific custom deployment procedure) , security audit (issues related to opening several ports: 80 and 8080)	On premises deployment, distributed deployment, no deployment guide outside the appliance, no containerization, distro outdated.
Non free technology inside the appliance	not a problem	not a problem	not a problem	yes	not a porblem	not a porblem	not a porblem	Based on previous experience with Allegrograph I fear that it's maintainance and integration might be very demanding.
Limitation with the current architecture and technologies	4store and Ruby (hard to find reference material when we met problem)	?	?	No possibility to switch the triple store, no dockerisation, good Ruby developers are rare and expensive	4store, everything in the triple store has its limit (considering using Neo4j for the mapping repository) Metadata model refactoring Modernize RoR and front end technos We need a bit of innovation.	RubyOnRails	Complexity makes troubleshooting, security standards compliance and maintenance difficult.	Deployment, old distro and visualisation libraries, comment system
Role	**Observer (Brainstorming, return of experience, share ideas + Content coordination + Outreach and communication)**	**Observer (Help and support + Training actions + Content coordination)**	**Contributor (Brainstorming, return of experience, share ideas, Training material and activities, Outreach and communication)**	**Contributor (Brainstorming, return of experience, share ideas, New technical (joint) developments, Outreach and communication, Training material and activities)**	**Contributor (New technical (joint) developments, Brainstorming, return of experience, share ideas, Organisation and management)**	**Contributor (Help and support, Training actions, Organisation and management)**	**Observer (Help and support + New technical (joint) developments)**	**Contributor (Brainstorming, return of experience, share ideas, New technical (joint) developments, Outreach and communication)**
Collaboration means	Regular general meetings + Topic specific meetings / followups + Open source code sharing practices	Topic specific meetings / followups + Collaborative documents (GDrive) + Mailing lists	Regular general meetings + Open source code sharing practices	Topic specific meetings / followups + Open source code sharing practices + Centralized project board	Regular general meetings + Centralized project board + Slack chats	Topic specific meetings / followups + Open source code sharing practices + slack chats	all	Regular general meetings, Topic specific meetings / followups, Open source code sharing
Inter-portal feature	✅	⛔	⛔ (not sure)	✅	✅	✅	✅	✅
Commun project funding	⛔ (hard for our team to get support for the Alliance)	✅ (very excited to submit a project together)	⛔ (not sure)	✅ (maybe the possibility of creating a non profit organisation mainly publically funded)	✅ (Eventually a partner spin-off or startup can be a plus in the ecosystem, EOSC)	✅ (very excited to submit a project together, Collaborative funded project)	⛔	✅ (Joining the Apache Software or Eclipse Foundation would create a solid structure and trust for Industrial Open Source Involvement)
Lead a shared action	?	Mappings	All	Scientific dissemination, explicit outreach	SKOS, mappings, metadata, policies, outreach	Ecology	Evolution towards SKOS management	Communication, Dissemination, Training, Workshops

OntoPortal philosophy

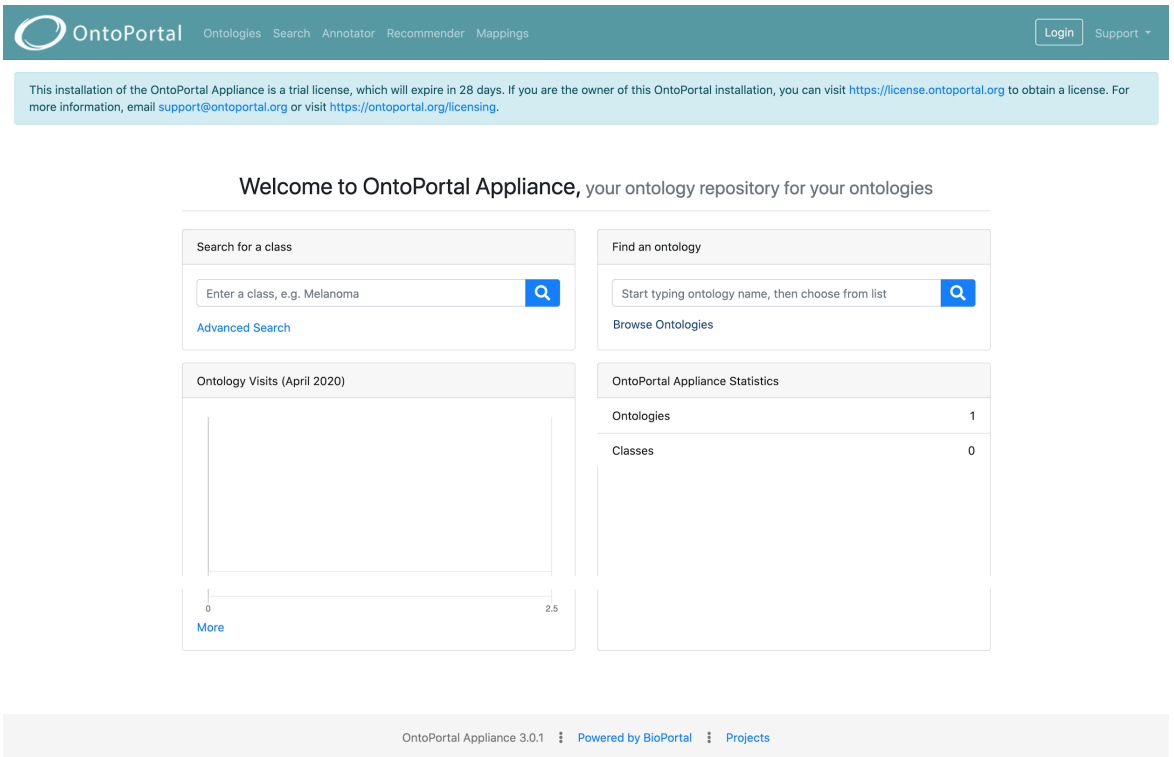
- OntoPortal installations do not own ontologies, they just serve them. Most portals have adopted an open approach with a flexible editorial policy, where users can upload resources themselves and the evaluation is left to the community. The OntoPortal project teams do not decide if an ontology is good or not.
- Semantic resources can be hosted in several installations, but a better coordination of the content is needed. The idea of a “scientific board” including semantic experts outside of the Alliance could be an approach.
- Motivation is to make use of ontologies easier for users who are not familiar with semantic technologies.
- OntoPortal becomes useful in the lifecycle of semantic resources at the point when releasing a specific version to archive or share it becomes necessary.
- The OntoPortal Alliance does not have a legal entity and it is likely to want one in the future.
- Some communication is needed including with scientific publications.
- OntoPortal technology is as modular as we can make it and contains the basic components that everyone wants. Portal specific customizations does not belong in the main code.
- We adopt an open source philosophy in which we believe that by investing efforts on the shared developing code for the shared project, we can progress more and more efficiently for our own project.
- We rely on standardized semantic web formats and technologies in order to develop generic, domain agnostic and compliant software that can be applied to a wide range of uses which may not be anticipated.



OntoPortal outreach and funding

People engaged in the discussions about outreach and funding recognized the need for improved focus in both areas. We prioritized the most important outreach assets and the meeting identified several strategies that could increase our opportunity for obtaining funding and visibility.

OntoPortal technology



Feature requests identified by the partners

Miscellaneous

- Implement a read-only version of OntoPortal i.e., an installation with no user contributions (ontologies, projects, mappings, notes, login).
- Creation of different roles.
- Support another open source triplestore backend (beyond AllegroGraph which is not free).
- SHACL shapes-based validation of semantic resources (what we need in OntoPortal, no lessons on how to develop an ontology).
- Refactor the ontology submission process so that a user is invited to come back after parsing to validate/curate the outputs of the process and eventually fix a few issues and edit metadata BEFORE full availability in the OntoPortal.
- Refactor the notes to be more modern and connected to the semantic resource developer practices.
- Consolidate all the documentation available (deployer guide, developer guide, user guide) and setup a system that allows to factorize documentation between OntoPortal installations.
- Refactor the ontoportal.org web site (https://ontoportal.org) and use it as a landing page for: the technology, the alliance and the documentation.
- Re-develop an OntoPortal plug-in for Protégé.
- Develop a link to an ontology editor specific to a semantic resource similar to what EcoPortal has done for VocBench (can be based on a metadata attribute)
- Better support of ontology imports and re-uses of objects (classes, properties, instances, concepts) from other semantic resources.
- Support a unique, solid and relevant term proposal process (currently OntoPortal has three: “Add proposal”, “Provisional classes” and OntoloBridge)

- Better support ontology diffs and improve perception of what items have changed between versions and overall ontology evolution.
- Multilingual support for ontology contents
- Multilingual support for the UI
- Select a shared classification to define and harmonize our categories cross portals

Metadata

- Re-align EcoPortal and AgroPortal metadata model.
- Merge the DOI feature of EcoPortal for portals with a rich metadata model (Agro/SIFR/Industry).
- Implement an incrementally augmentable minimum metadata model allowing any portal to incorporate new metadata properties 'one-by-one'.
- Enable a parameterizable metadata model.
- Refactor the Summary page for portals with a rich metadata model (Agro/SIFR/Eco/Industry).
- Whatever the metadata model, consolidate the automatic metadata extraction and generation to ease the management/curation of metadata.
- Whatever the metadata model, find a way to import metadata from external sources e.g., a VOAF file or OBO Foundry YAML description.
- Better description of OntoPortal page (not only Summary) with Schema.org for better indexing by Google.
- Implement a metadata search service in each portal and cross portal.
- 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).

Mappings

- Full support of SSSOM for the mapping repository.
- 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
- Implement better views for mappings including an evaluation view and an overview of mappings between any n ontologies
- Support bulk upload and download of mappings
- Make existing mappings selectable/viewable by their origin (e.g, show only API mappings, or only automated mappings).

Federated access

- 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. (Can be extended to a special section for commercial members, for a small donation.)
- Avoid repeating authoring of metadata from one portal to another.
- Implement a federated search of content. Grab either statistics (number of hits) or even results (to produce snippets) in the querying interface.
- Federated Recommender. When a query is passed to the recommender, select the option to query also the other Recommenders.
- Implement a cross portal metadata search service that will facilitate the identification of the semantic resources.

DevOps and environment

- Docker container based setup/installation of OntoPortal
- Implement and describe an easily deployable environment test
- Regression testing automatically on GitHub code updates
- A configurable OntoPortal appliance
- Add monitoring tools
- Create a SPARQL query editor/viewer. Minimally plug-in the viewer of the underlying used triple store (4store and AllegroGraph have one).
- 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.
- Implement data (ontology) migration scripts
- Better support versioning of code and identification for which code is currently running on which appliance
- Address the end of support of CentOS
- Move to a more generic (and GDPR complaint) tool for analytics beyond GoogleAnalytics.
- Create an option to enable deployment of a separate public SPARQL triplestore.

Feature requests identified during the public track

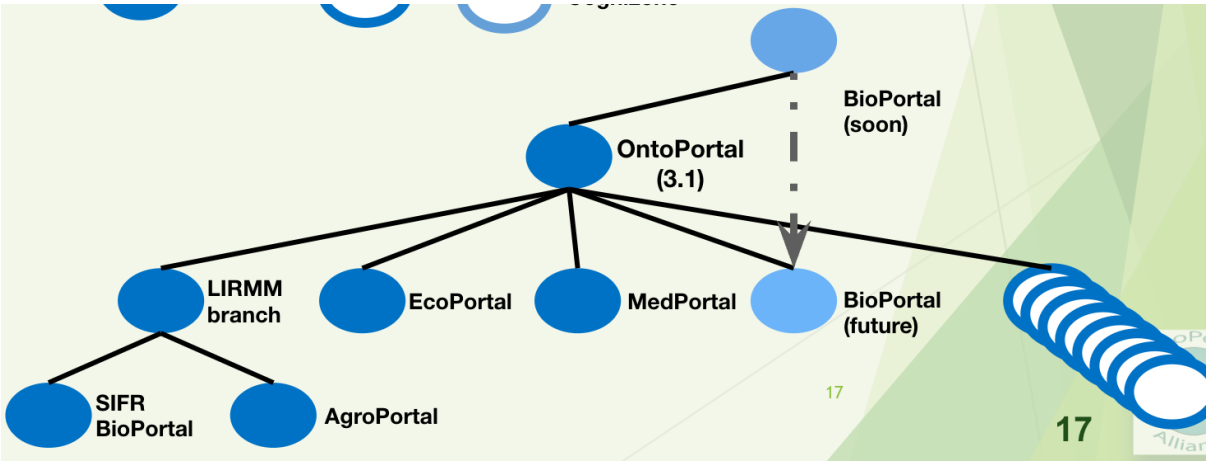
- Docker container based setup/installation of OntoPortal.
- Ultimately OntoPortal could be installed “pushing one button” within the appropriate research infrastructure (EOSC?)
- Better support of SKOS (something more aligned to what SKOSMOS does)
- Better integration with semantic resource editors especially VocBench and Protégé
- Better coordination when a semantic resource is hosted in multiple repositories in the alliance
- Should be easy for final users to find contents on different OntoPortal installations i.e., they must be connected somehow
- History view for the evolution of a semantic resource (historical influences between ontologies view as a graph)
- Semantic validation functionality that will automatically check a semantic resource with respect to identified guidelines (plug-in architecture)
- Better capture semantic resource status (manage obsolete resources)
- Better handling of ontologies that are knowledge bases (i.e., OWL model with many instance data)
- Link OntoPortal semantic resources to external datasets using or annotated by the semantic resources
- Multilingual support
    - Multilingual support for ontology contents?
    - Multilingual support for the UI?
- Better support of the Annotator as many many data are still in text form
- Better documentation is needed (at very level) maybe a series of webinar or training materials
- Video tutorial for the installation of OntoPortal
- Tutorial on how to setup a development environment for a new OntoPortal developer
- Keep a full open source distribution of the OntoPortal Appliance
- The portal need to facilitate ontology competency questions e.g., by showing the type of questions/queries an ontologie allows to answer.

Decisions & next actions

1. Any new functional feature developed by OntoPortal developers should be designed as a new module that can be parameterized in the config files and activated (or not).

2. From now on the Alliance meetings will follow this structure:
    a. OntoPortal governance monthly meeting (2nd Thursday of each month at 1500 GMT). Organized by J. Graybeal (deputy C. Jonquet). Invites will be sent out to the OntoPortal Alliance management mailing list ontoportal-alliance@lists.stanford.edu
    b. Topic-specific meetings organized by a topic leader on a specific pre-selected topic. Invite will be sent out to both management and developer mailing list ontoportal-developers@lists.stanford.edu

3. From now on we must work to move to a centralized repository on GitHub https://github.com/ontoportal/ and work as an "open source" project on this repo.
    a. Beside the code repository and readme files, the issues, teams, discussions, and actions mechanisms of GitHub will be used.
    b. Alliance partners shall "fork" the code repositories to enable traceability and collaborative contributions via pull requests.
    c. BioPortal will have to make a list of the tasks required to make BioPortal a fork of the OntoPortal repo, rather than a master.
    d. A system must be developed by which pull requests to the OntoPortal repo can be efficiently evaluated and accepted (or not) by the OntoPortal community.



4. We should reproduce such workshop at minimum yearly. Maybe if the motivation and context is here the next one should be before right before summer 2023. Stanford seems a logical location for the next meeting but might cost more to the Alliance. (Possibly we could find another location not too far from Stanford, e.g., Santa Cruz.) We also need an organizer or co-organizers.

Topics and topic leaders

Over the next months, in addition of the actions and collaborations through GitHub, we will organize the development and actions related to specific discussed and identified topics. The topic leaders will be kicking off the required meeting and collaboration means for their specific topic. No need to have everyone in the Alliance represented to move these topics forward.

- Documentation [Xeni Kechagioglou]
- SKOS support [Clement Jonquet]

- UI/UX refactoring [Syphax Bouazzouni]
- Working with a centralized GitHub repository [Jennifer Vendetti]
- Metadata [Clement Jonquet & Ilaria Rosati]
- Mappings & SSSOM [Arkopaul Sarkar]
- Automatic deployment procedures [Alex Skrenchuk]
- Communication and publication [Hedi Karray]
- Federated services [Naoel Karam]
- Term proposal [Xeni Kechagioglou & Jennifer Vendetti]
- Ontology versioning and diffs [Ralph Schäfermeier]
- AllegroGraph transition [Misha Dorf]
- New markets for OntoPortal [Agis Papantoniou]

Technical points and bugs

We are listing here the list of issues identified and that will need to be moved to GitHub in time.

Ontology parsing

- 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
- 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
- Code cleaning
- Real time notifications of processing status
- Optimizing process for large ontologies
- Implement notifications so re-using ontologies/users will know when an ontology changes
- Implement a better ontology comparison tool (currently Bubastis). Facilitate the vision/ understanding of what has changed.
- Connect to tools to Vidoco to produce HTML doc for (small) semantic resources

Triplestore evolution

- Finalize support and transition to AllegroGraph
- 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 distribution with Virtuoso
- Implement a triple data migration script to move data from a 4store to another triplestore

Front end change

- 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 cache after ontology submission for every other user to see the new submission
- Enable to see a Summary page change (metadata edition) right away
- Update Bootstrap version
- Eliminate the old JS libraries
- Update our JS so we aren't forced to use jquery-migrate
- Refresh the widgets
- What would be the future of notes/proposals

- Refactor multiple AJAX calls
- Upgrading to Rails 7 (done on ontoportal-lirmm)
- Adopt Stimulus and Hotwire to avoid JavaScript (done on ontoportal-lirmm)
- Adopt TurboFrame (done on ontoportal-lirmm)
- Adopt ViewComponents as reusable view components (done on ontoportal-lirmm)
- Migrate and define all our reusable ui component as view components
- Use Lookbook framework for previewing/testing the UI components
- Rethink the UI and user experience
- Make the UI easily and more customizable

Observations

- We covered as many topics as possible. List of uncovered items for future meetings.
- We did not reach out to a specific category of stakeholder: the persons who have installed the OntoPortal Appliance but have not used it to make a public ontology repository. Stanford has a list (pre-license time and post-license time).
- Semantic annotation ("data labeling") is still a very important use case of semantic resources especially with the needs of supervised machine learning. Eventually we might need to come back to this topic.
- Many people would have liked to watch the recording of the public sessions.
- The momentum has been described several times with opportunities showing up and time changing especially with the large adoption of Open Science. We have a card to play these days because with Open Science and the FAIR principles, people realize they don't know how to do it.