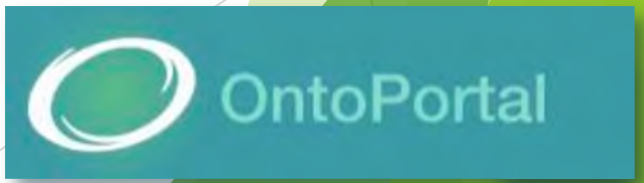
BioPortal to OntoPortal to OntoPortal Alliance: Meeting Needs

Presentation for NKOS 2023

John Graybeal

Technical Program Manager for BioPortal 2016–2023

OntoPortal Alliance Coordinator 2017-2022



- BioPortal
- OntoPortal
- OntoPortal Alliance

bit.ly/bp-op-opa

The Principals

Mark Musen, M.D., Ph.D.

Stanford University School of Medicine

Stanford Center for Biomedical Informatics Research (BMIR) (Director)

Principal Investigator of BioPortal

https://med.stanford.edu/profiles/Mark_Musen

The Principals

Clement Jonquet, Ph.D.

Senior Researcher, INRAE &

Associate Researcher, University of Montpellier

Principal Investigator, AgroPortal

https://jonquet.mystrikingly.com/

Need: Common biomedical concepts repo
Result: National Center for Biomedical Ontology (2004)

- Funded development of repository for biomedical ontologies
- BioPortal launched in 2006
  - Educating community about the repository and its importance
  - Accumulated an average of 50 or so ontologies per year
  - Allowed anyone to submit an ontology for public or private use
  - Shared backend service software with providers like Marine Metadata Interoperability Ontology Registry & Repository
- Significant funding until about 2016

BioPortal Project today

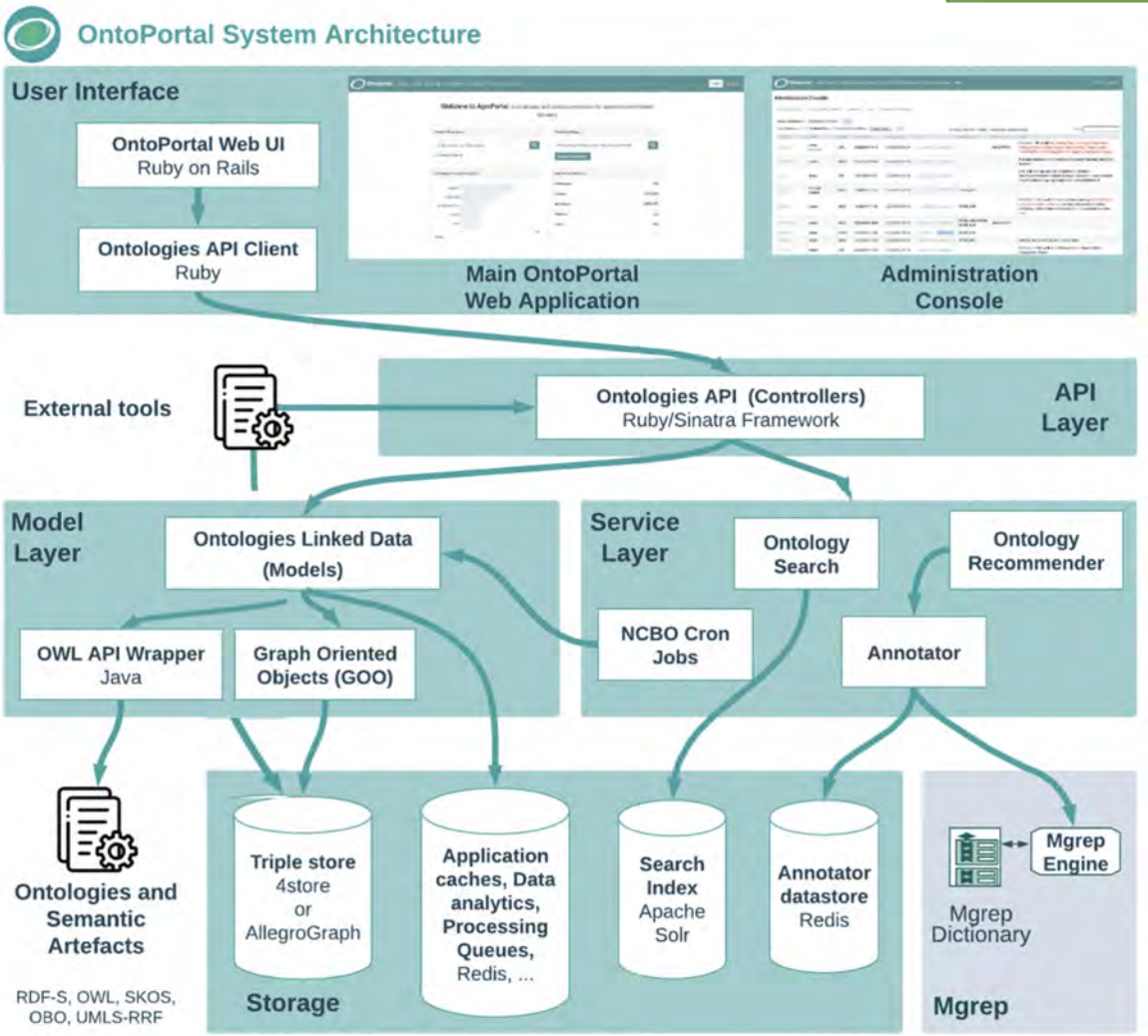
- Parent Project: Stanford Center for Biomedical Informatics Research
- Portal Purpose: Biomedical and other community ontologies
- Users: >16,750
  - Ontologies >>

# Ontologies	Public	Private	TOTAL
Complete	1082	337	1419
Partial Views	183	22	205
TOTAL	1265	359	1624

Need: Community repository services
Result: NCBO Virtual Appliance (2012)

- Always access the source code in GitHub *ncbo* repo
- Easily deploy your own copy of BioPortal repository software
- No cost to deploy privately, publicly, or commercially
- BioPortal team support: support@bioontology.org
- Includes most BioPortal services
- Released cloud-based version in 2016(?) as OntoPortal Appliance Amazon Machine Instance

OntoPortal Appliance: Architecture Today



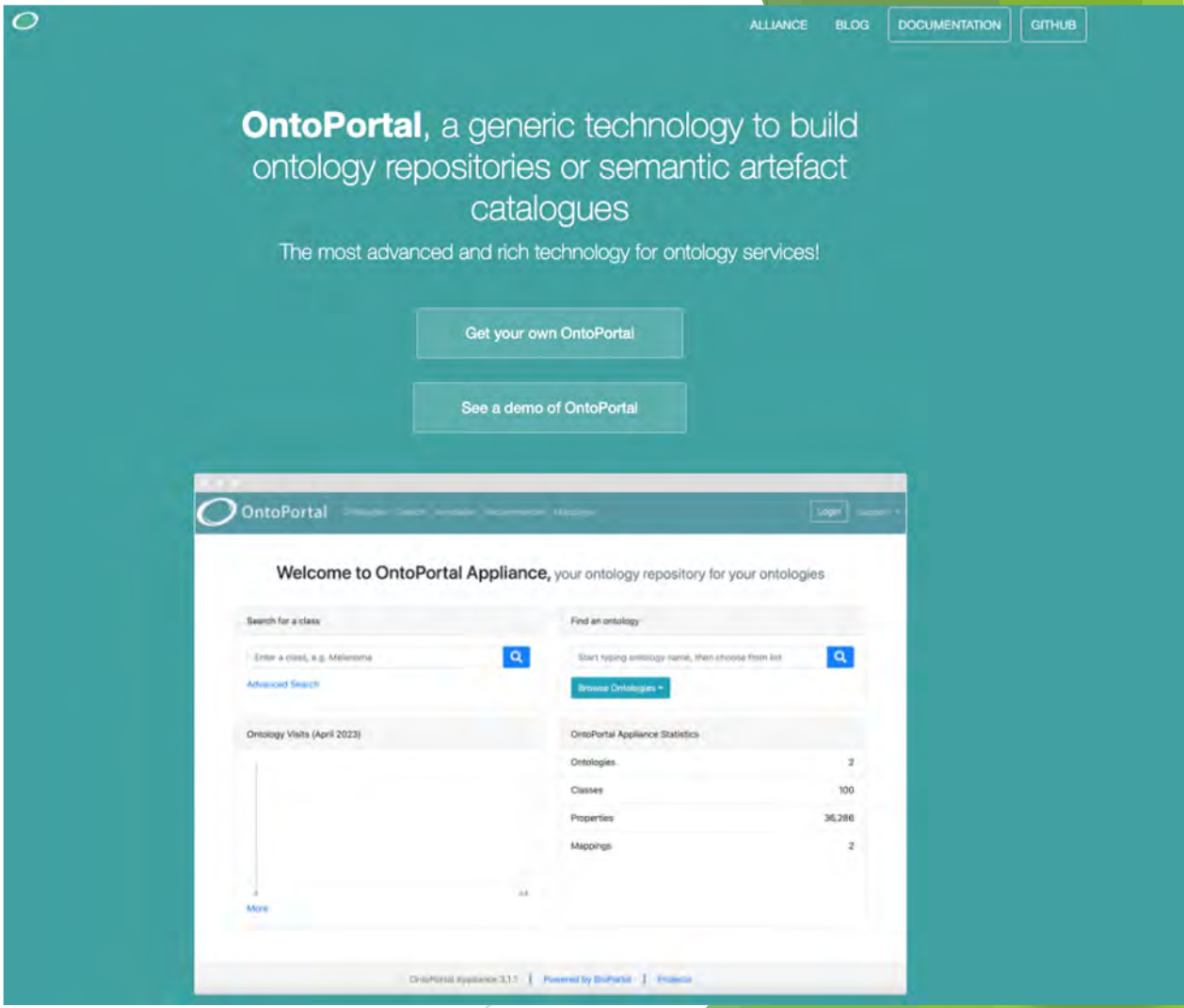
Need: Store non-biomedical ontologies too
Result: BioPortal supports any domain content (~2016)

- Noble reasons
  - Science often requires terms from many domains
  - Difficult to draw precise lines
  - Provide more services to more communities
- Practical reasons
  - BioPortal doesn't actively curate ontologies
  - More general-purpose tools like CEDAR used BioPortal

Need: Support like-minded communities
Result: Created the OntoPortal Alliance (2018)

- Noble reasons
  - Provide support for each other's work
  - Advance our common code base faster
  - Present a unified 'community of interest'
- Practical reasons
  - Identify and understand shared requirements
  - Increase 'market share' in semantic community
  - Increase the level of code sharing (not just adoption)
  - Get to reuse other people's work

ontoportal.org



Be part of a **community**

BioPortal
The world's most comprehensive repository of biomedical ontologies

SIFR BioPortal
A repository for French biomedical terminologies and ontologies

AgroPortal
A vocabulary and ontology repository for agronomy and related domains

EcoPortal
The LifeWatch ERIC repository of semantic resources for ecology and related domains

MedPortal
A repository for Chinese biomedical terminologies and ontologies

MatPortal
The ontology repository for materials science

IndustryPortal
A common ontology portal for industry and related domains

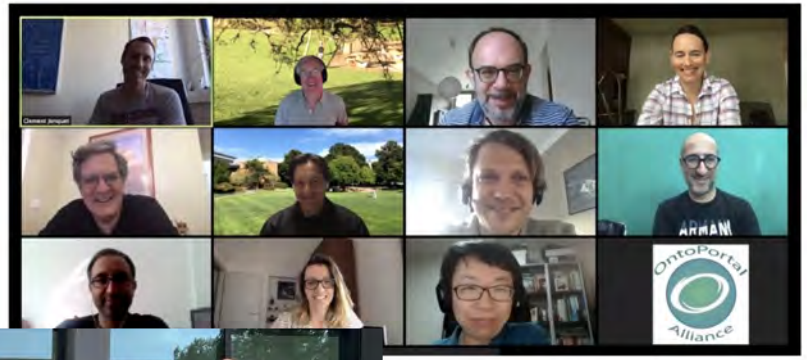
EarthPortal
A semantic artefact repository dedicated to Earth sciences

BiodivPortal
A semantic artefact repository for biodiversity

Repos and Workshops

- [ ] **AgroPortal**
  Member of the AgroPortal and SIFR BioPortal team mostly at LIRMM and MISTEA
- [ ] **BiodivPortal**
  NFDI4biodiv team working on a dedicated OntoPortal
- [ ] **BioPortal**
  Members of the BioPortal team mostly at Stanford BMIR.
- [ ] **CogniZone**
  Member of the Cogni.zone SME team.
- [ ] **EarthPortal**
  Members of the EarthPortal team mostly at CNRS and DataTerra
- [ ] **EcoPortal**
  Members of the EcoPortal team mostly at LifeWatch ERIC
- [ ] **IndustryPortal**
  Members of IndustryPortal team mostly at ENIT
- [ ] **MatPortal**
  Members of the MatPortal team mostly at Fraunhofer
- [ ] **MedPortal**
  Members of the MedPortal team mostly at BMICC.

2020

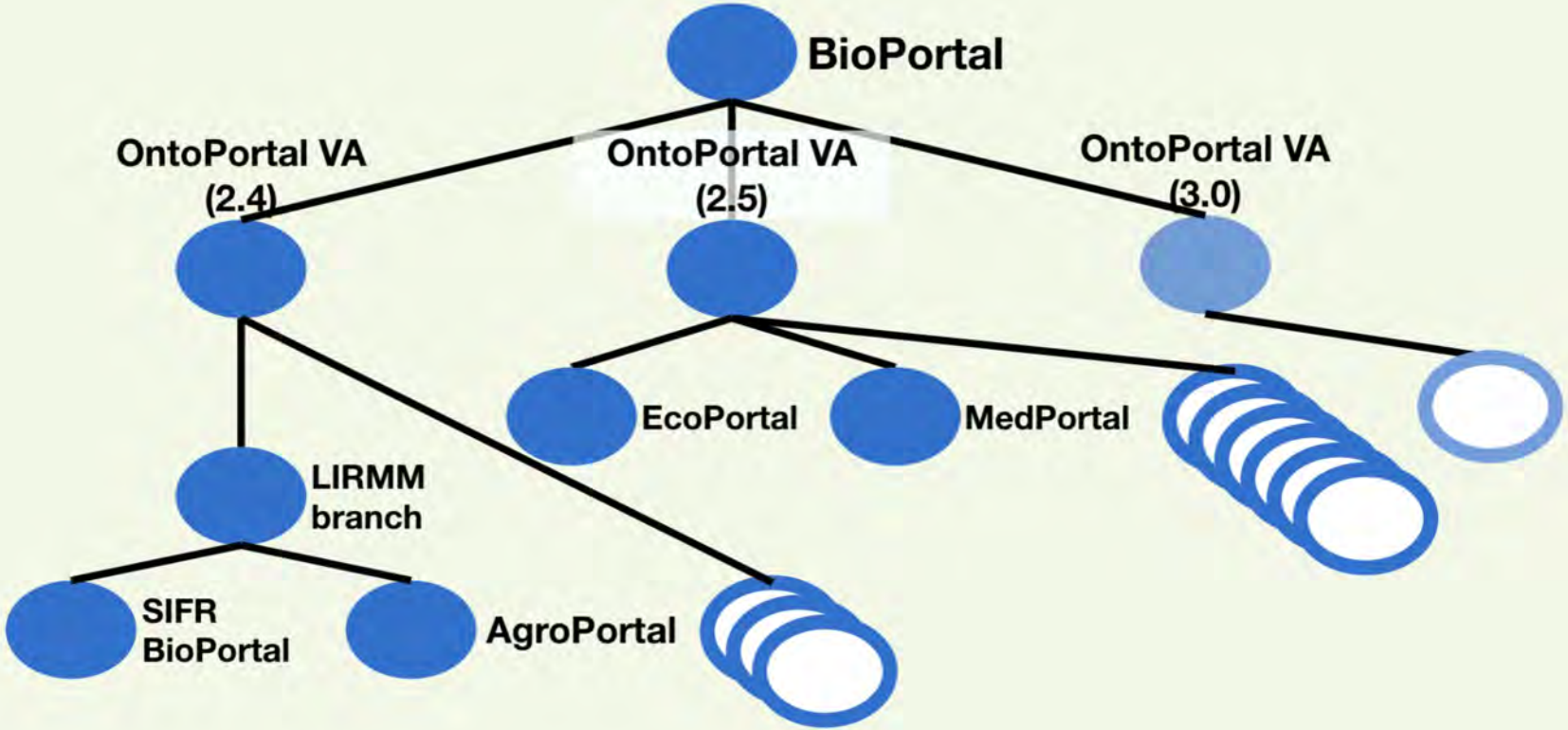


2022

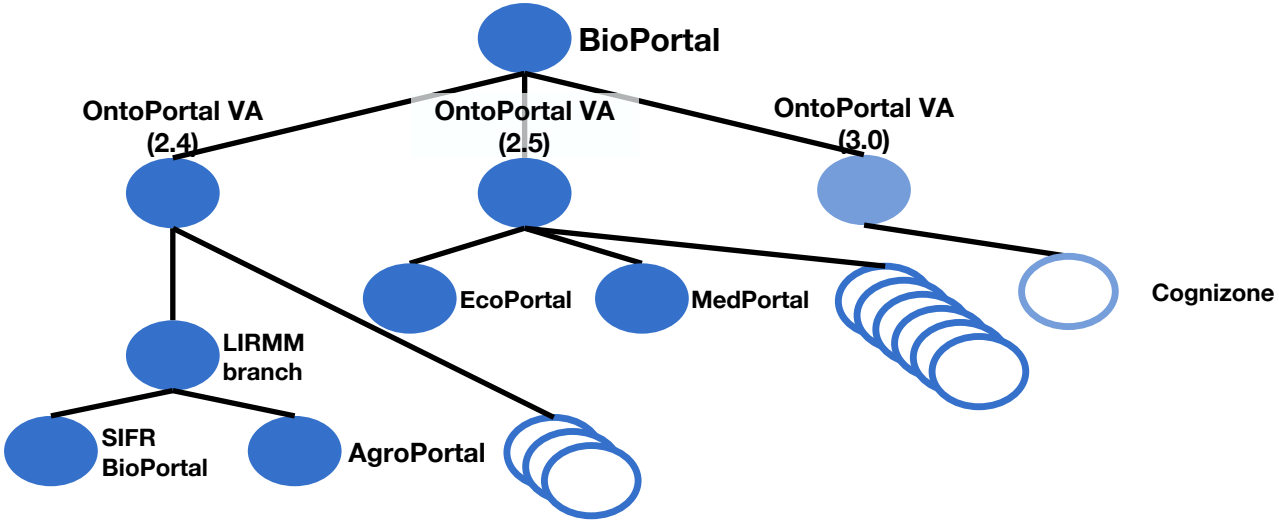
2023

How do we keep deployments aligned?

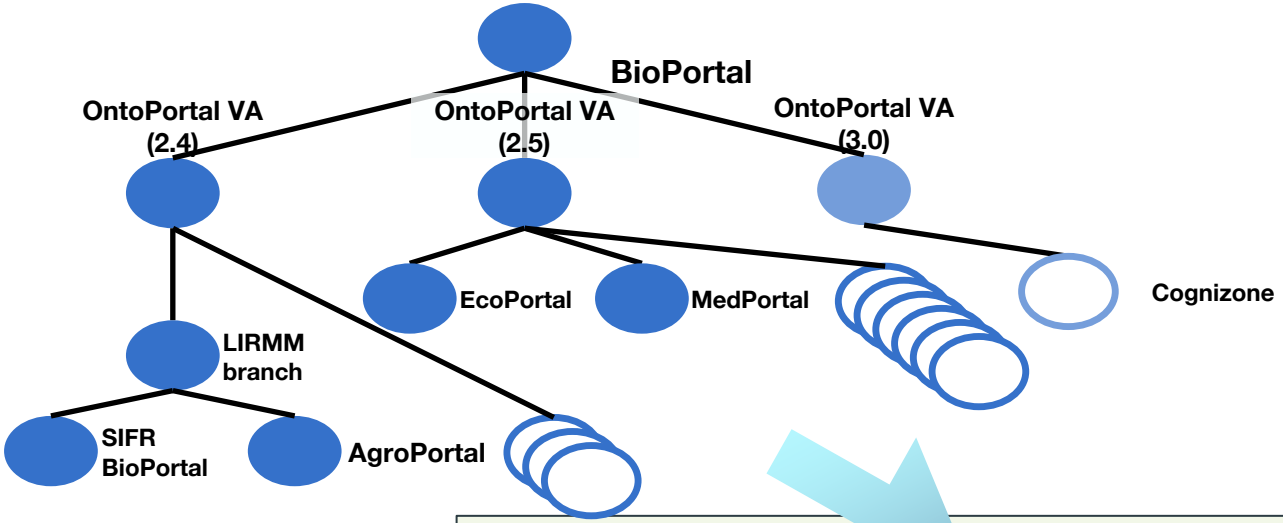
(Our starting point, years ago…)



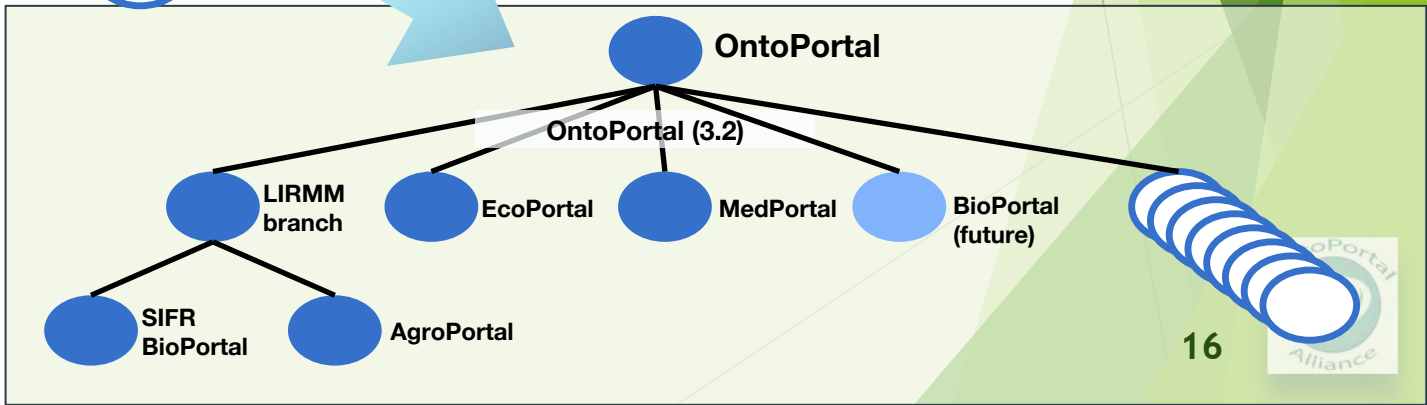
How do we keep deployments aligned?



How to transition to what we really want?



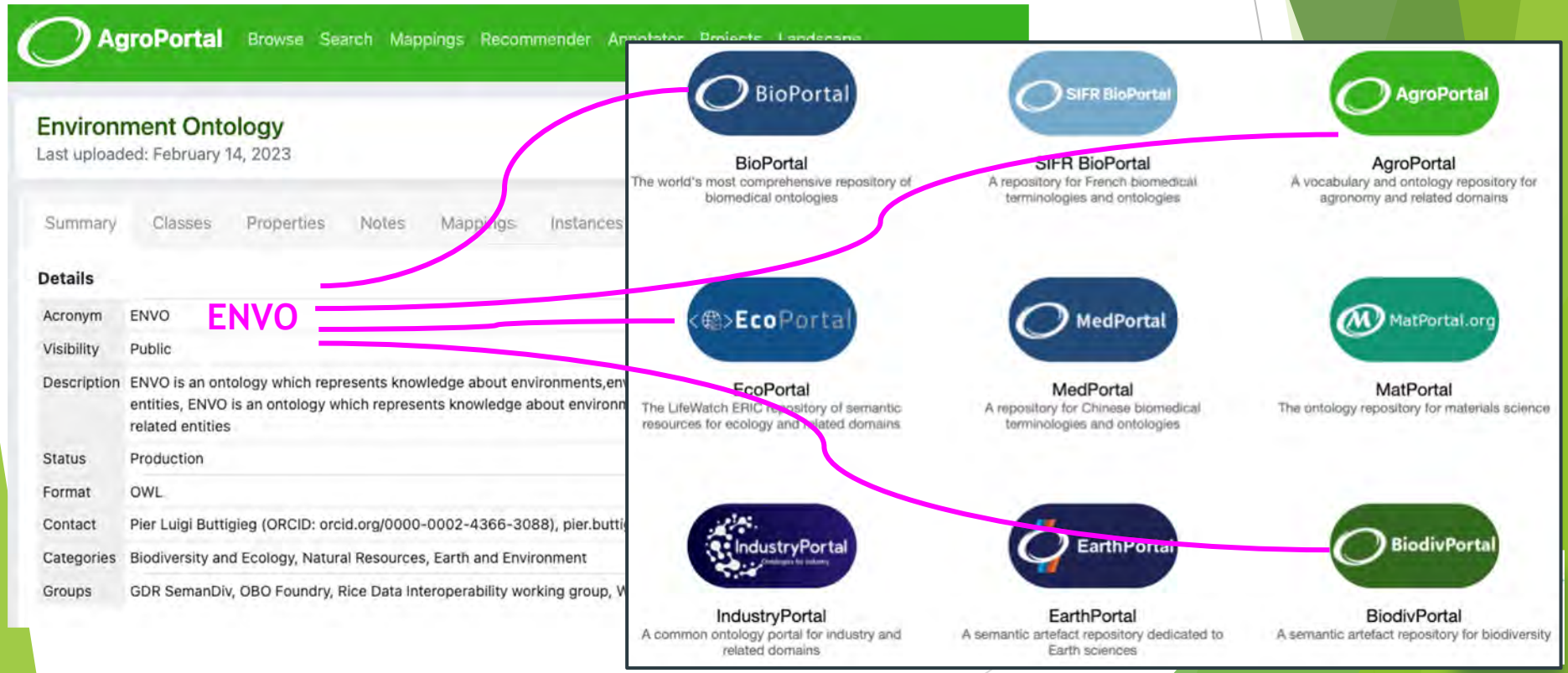
Many steps needed with mutual Patience and support



Need: Common code distribution model
Result: New GitHub distribution model (mostly)

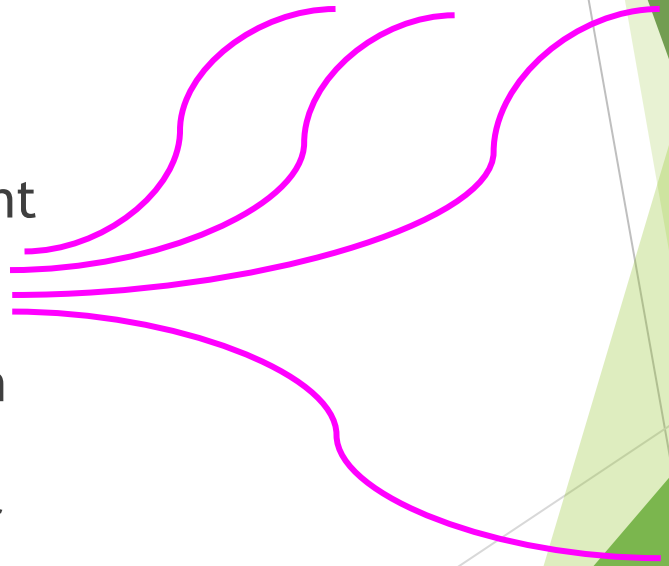
- “Who’s on top?” in this diagram matters, some. And:
  - Where are most changes originating?
  - How many changes are there?
  - Who has time to evaluate and integrate changes?
- Behind the scenes best practices are essential
  - Good shared coding practices
  - Good separation of concerns/modularization
  - Really crisp regression testing
- Once all in place, sharing code will be much easier

Need: Federation agreement for ontology sharing



Need: Federation agreement for ontology sharing
Result: Agreed ‘ontology federation’ practices

- Is one of these primary?
- How are copied kept current?
- What if a repo needs a different version?



We have agreement in principle on how to manage this kind of shared resource across all the Members of the Alliance.

Need: Diverse features for different projects

BioPortal: A lot of infrastructure work improving these…

- Submission/Ingest
  - Format Conversions
  - Versioning
  - Differencing
  - Metrics re Vocab.
  - Interpretation/ Property Mapping
- Search
  - Ontology Name
  - Terms
  - Best Terms
  - Attribute/Value
  - Metadata
  - Triples
- Subselection
  - Views (partial voc.)
  - Slices (partial repo)
  - Preferred Vocabs

- Terms
  - Browsing
  - Create ID
  - Resolve ID
  - Term Pages
  - Mappings
    - Auto–syntax
    - Auto–identifiers
    - Manual–user
- Vocab. Recommend
  - Keyword-based
  - Text-based
  - Multi-vocabulary
  - Controllable criteria
  - Selectable vocabs
- Annotation
  - Keyword/text based
  - Selectable vocabs

Widgets
- Term completion
- Term browsing
- Visualization

Social
- Reviews
- Notes/Comments
  - Ontology
  - Term

Usage Tracking

Visualization
Licensing

- # API requests

API Features
Resource Index
Visualization

Need: Diverse features for different projects
Result: Agile methods to share/test features

AgroPortal major features: multilingual ontologies support; complete SKOS capabilities; in-app FAIR evaluation of ontologies

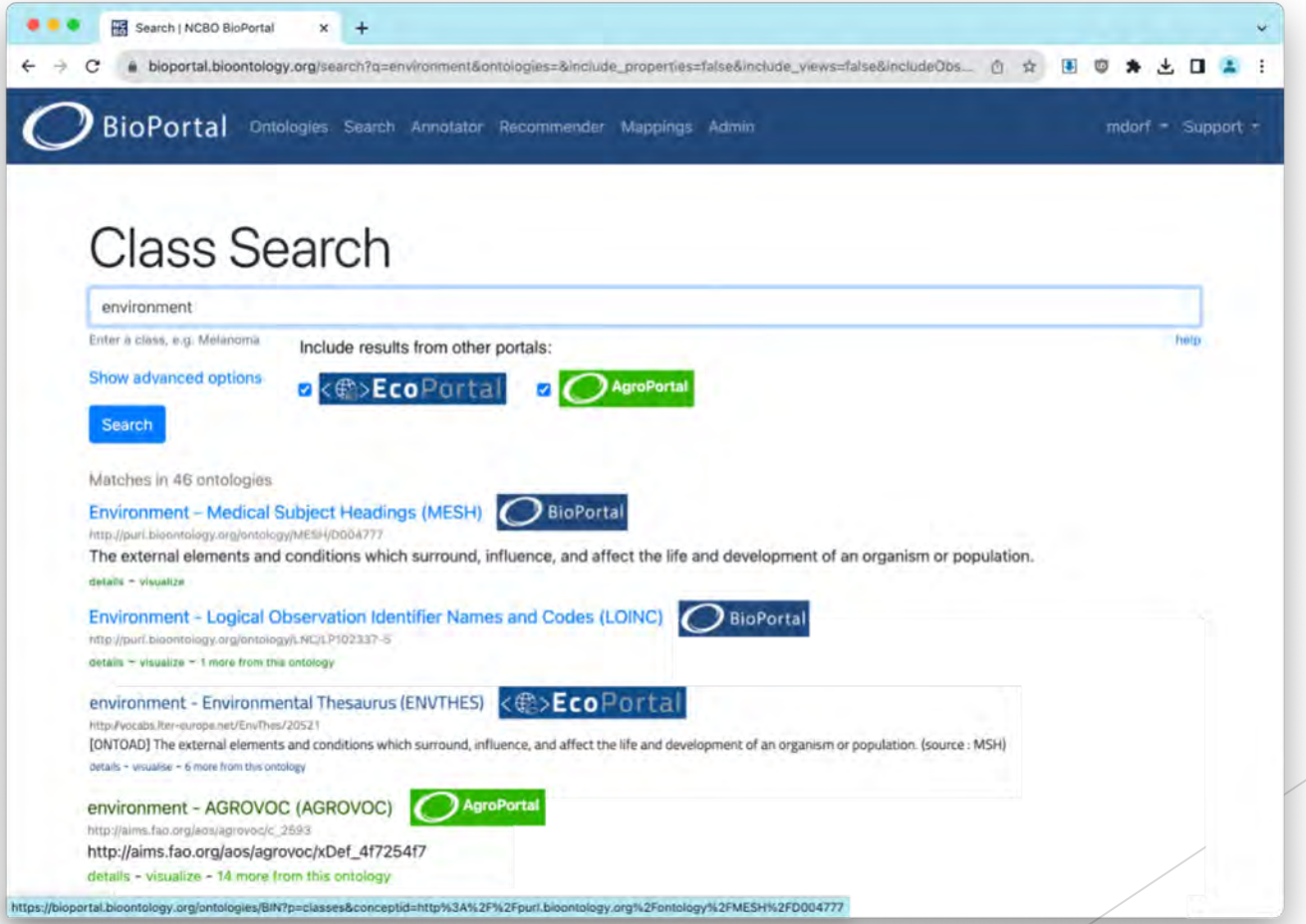
AgroPortal other features: new administrations tools (metadata, groups and categories); some new UIs; some UI internationalization; consolidated metadata model

BiodivPortal major features: diff and versioning support; SSSOM (mapping) support; iADOPT integration

MedPortal major features: replace Google page tracking; Chinese internationalization; CSV imports

EcoPortal major features: VocBench and ShowVoc integration; assigning DOIs; Keycloak-based Single Sign-On (SSO)

Need: Federation of features (e.g., search)

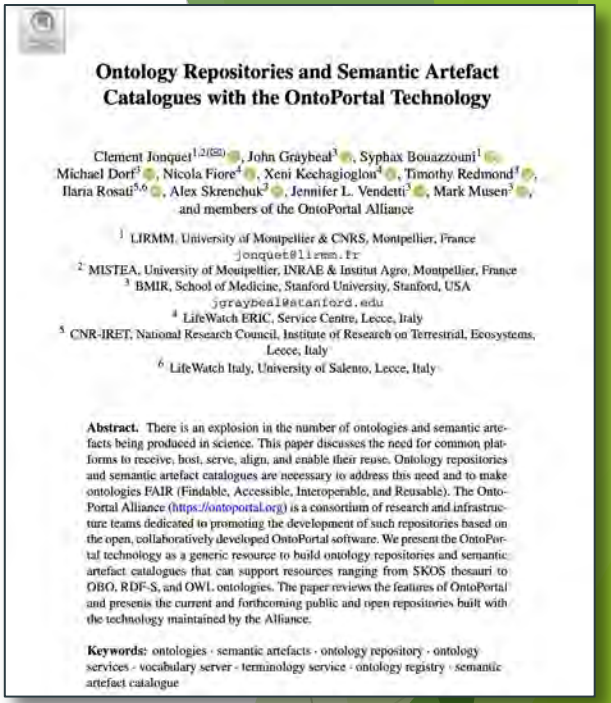


Need: Federation of features (e.g., search)
Result: Initial strategies for federating search

- Agreement that search is first feature to federate
- Agreement on what federated search is:
  - While searching on a term, user can choose to extend search to other OntoPortal instances
- Agreement on what is found:
  - Matching entities, sorted in some way
  - Further analysis needed of best sorting approach
- Ideas matured in other areas
  - How to best display results
  - Best architecture for responsiveness & reliability

Need: A publication
Result: A publication!

- Clement Jonquet, John Graybeal, Syphax Bouazzouni, Michael Dorf, Nicola Fiore, et al.. **Ontology Repositories and Semantic Artefact Catalogues with the OntoPortal Technology**. ISWC 2023 - 22nd International Semantic Web Conference, Nov 2023, Athens, Greece.



Ontology Repositories and Semantic Artefact Catalogues with the OntoPortal Technology

Clement Jonquet[1,2(✉)], John Graybeal[3], Syphax Bouazzouni[1], Michael Dorf[3], Nicola Fiore[4], Xeni Kechagioglou[4], Timothy Redmond[3], Ilaria Rosati[5,6], Alex Skrenchuk[3], Jennifer L. Vendetti[3], Mark Musen[3], and members of the OntoPortal Alliance

[1] LIRMM, University of Montpellier & CNRS, Montpellier, France
jonquet@lirmm.fr
[2] MISTEA, University of Montpellier, INRAE & Institut Agro, Montpellier, France
[3] BMIR, School of Medicine, Stanford University, Stanford, USA
jgraybeal@stanford.edu
[4] LifeWatch ERIC, Service Centre, Lecce, Italy
[5] CNR-IRET, National Research Council, Institute of Research on Terrestrial Ecosystems, Lecce, Italy
[6] LifeWatch Italy, University of Salento, Lecce, Italy

Abstract. There is an explosion in the number of ontologies and semantic artefacts being produced in science. This paper discusses the need for common platforms to receive, host, serve, align, and enable their reuse. Ontology repositories and semantic artefact catalogues are necessary to address this need and to make ontologies FAIR (Findable, Accessible, Interoperable, and Reusable). The OntoPortal Alliance (https://ontoportal.org) is a consortium of research and infrastructure teams dedicated to promoting the development of such repositories based on the open, collaboratively developed OntoPortal software. We present the OntoPortal technology as a generic resource to build ontology repositories and semantic artefact catalogues that can support resources ranging from SKOS thesauri to OBO, RDF-S, and OWL ontologies. The paper reviews the features of OntoPortal and presents the current and forthcoming public and open repositories built with the technology maintained by the Alliance.

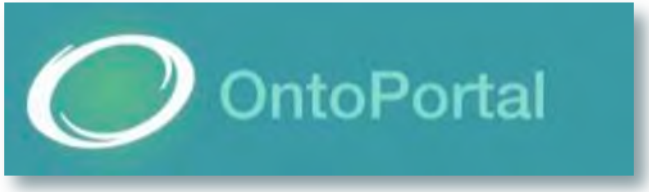
Keywords: ontologies · semantic artefacts · ontology repository · ontology services · vocabulary server · terminology service · ontology registry · semantic artefact catalogue

https://doi.org/10.1007/978-3-031-47243-5_3

In Conclusion

- BioPortal evolved as a feature-rich research platform
- OntoPortal provided the chance for a community to gather around the platform, all using their own funding models
- The OntoPortal Alliance is advancing on two fronts
  - Evolving the OntoPortal platform into a shared, production-level software product
  - Establishing a community capable of supporting and advancing the shared software vision
- The whole platform—code, documentation, web presence—is increasingly solid, supported, and sustained.

Thank You to the Teams



OntoPortal Alliance: Synchronizing and mutualizing research and development efforts