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 >>

<table>
<thead>
<tr>
<th># Ontologies</th>
<th>Public</th>
<th>Private</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>1082</td>
<td>337</td>
<td>1419</td>
</tr>
<tr>
<td>Partial Views</td>
<td>183</td>
<td>22</td>
<td>205</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1265</td>
<td>359</td>
<td>1624</td>
</tr>
</tbody>
</table>
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, a generic technology to build ontology repositories or semantic artefact catalogues

The most advanced and rich technology for ontology services!

Get your own OntoPortal

See a demo of OntoPortal

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 the IndustryPortal team mostly at ENIT

- **MatPortal**
  Members of the MatPortal team mostly at Fraunhofer

- **MedPortal**
  Members of the MedPortal team mostly at BMICC.
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

ENVO
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...

<table>
<thead>
<tr>
<th>Submission/Ingest</th>
<th>Terms</th>
<th>Widgets</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Format Conversions</td>
<td>• Browsing</td>
<td>• Term completion</td>
</tr>
<tr>
<td>• Versioning</td>
<td>• Create ID</td>
<td>• Term browsing</td>
</tr>
<tr>
<td>• Differencing</td>
<td>• Resolve ID</td>
<td>• Visualization</td>
</tr>
<tr>
<td>• Metrics re Vocab.</td>
<td>• Term Pages</td>
<td>• Social</td>
</tr>
<tr>
<td>• Interpretation/Property Mapping</td>
<td>• Mappings</td>
<td>• Reviews</td>
</tr>
<tr>
<td>• Search</td>
<td>• Auto-syntax</td>
<td>• Notes/Comments</td>
</tr>
<tr>
<td>• Ontology Name</td>
<td>• Auto-identifiers</td>
<td>• Ontology</td>
</tr>
<tr>
<td>• Terms</td>
<td>• Manual-user</td>
<td>• Term</td>
</tr>
<tr>
<td>• Best Terms</td>
<td></td>
<td>Usage Tracking</td>
</tr>
<tr>
<td>• Attribute/Value</td>
<td></td>
<td>Visualization</td>
</tr>
<tr>
<td>• Metadata</td>
<td></td>
<td>Licensing</td>
</tr>
<tr>
<td>• Triples</td>
<td></td>
<td>• # API requests</td>
</tr>
<tr>
<td>• Subselection</td>
<td></td>
<td>API Features</td>
</tr>
<tr>
<td>• Views (partial voc.)</td>
<td></td>
<td>Resource Index</td>
</tr>
<tr>
<td>• Slices (partial repo)</td>
<td></td>
<td>Visualization</td>
</tr>
<tr>
<td>• Preferred Vocabs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• Annotation
- Keyword/text based
- Selectable vocabs
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!


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

- LifeWatch ERIC
  - https://ecoportal.lifewatch.eu

- MedPortal
  - Peking Union Medical College
  - biomedicine (Chinese)
  - http://medportal.bmic.cn

- EcoPortal

- AgroPortal
  - INRAE, University of Montpellier
  - agri-food

- BioPortal
  - Stanford University
  - biomedicine
  - https://bioportal.bioontology.org

- SIFR BioPortal
  - CNRS, University of Montpellier
  - biomedicine (French)

- MatPortal.org

- IndustryPortal

- ENIT / OntoCommons
  - industry

- OntoPortal, a generic technology to build ontology repositories or semantic artefact catalogues
  - The most advanced and rich technology for ontology services!

- Welcome to OntoPortal Appliance, your ontology repository for your ontologies