Contribution ID: **122**   Type: **Presentation**

OntoPortal – An Open Technology for Discipline-Specific Terminology Repositories

Tuesday 14 October 2025 16:22 (11 minutes)

Background and Motivation

Controlled vocabularies and ontologies are essential for enabling data interoperability, discovery, and integration across domains. Repositories that host and expose these artifacts play a critical role in implementing the FAIR (Findable, Accessible, Interoperable, Reusable) principles. The OntoPortal Alliance, a collaboration of academic and commercial partners, supports the development and deployment of terminology repositories through a shared open-source platform. Originally derived from the National Center for Biomedical Ontology's BioPortal resource (https://bioportal.bioontology.org), OntoPortal has evolved into a modular, extensible, domain-independent solution for managing collections of discipline-specific controlled terminologies. The system has been adopted by diverse research groups internationally to archive and manage terminologies in biomedicine, agronomy, ecology, biodiversity, earth science, materials science, and other domains. The groups that have deployed OntoPortal to provide terminology services in these disciplines form a network of linked repositories, collaborating as part of an affiliation called the *OntoPortal Alliance*.

Despite the increasing importance of ontologies and controlled terminologies in implementing FAIR data practices, many communities lack sustainable, adaptable infrastructure to support the lifecycle management of such resources. Most existing solutions are either overfitted to other research areas, not reusable, or fail to interoperate effectively. OntoPortal addresses this gap by offering a unified, open-source foundation for building and federating terminology repositories across diverse scientific disciplines.

System Overview

OntoPortal provides a full-featured repository infrastructure for publishing, discovering, and interacting with ontologies—both interactively and programmatically (Figure 1). The system supports uploading and curating terminologies and ontologies, searching across the uploaded content, browsing terms, and annotating data using terminology-based services. OntoPortal handles multiple representation formats (viz., OWL, SKOS, RDF, OBO) and it provides multilingual support, metadata enrichment, version tracking, mappings among terms, and usage metrics. The platform exposes RESTful APIs and a web-based interface for both end users and terminology developers.
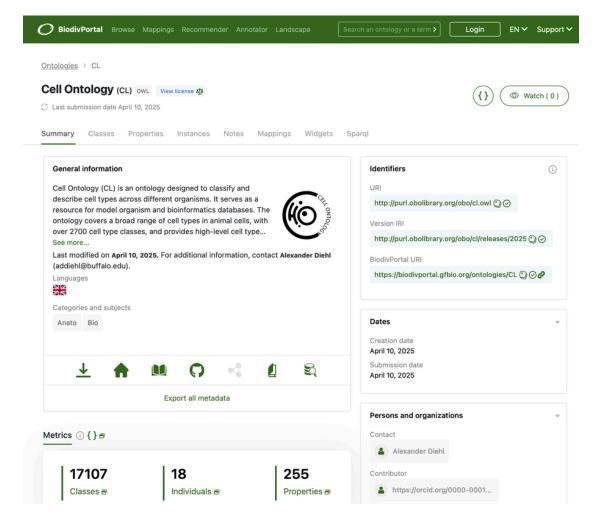
Figure 1: Figure1

Figure 1: Landing page of the Cell Ontology in the BiodivPortal implementation of OntoPortal.

OntoPortal's architecture includes backend services supporting RDF triple-store integration and indexing, and frontend components for user interaction. Management of terminology metadata is a central feature, and OntoPortal supports configurable metadata schemas including MOD (Metadata for Ontology Description). Term mappings, annotation services, change tracking, and user-generated content are integral to the system. Each deployment can adopt a flexible editorial policy, ranging from community-contributed catalogues to curated domain-specific repositories.

Recent work by OntoPortal Alliance members led to a major milestone: a federated release between AgroPortal, EcoPortal, EarthPortal, and BiodivPortal. This release will feature shared browsing across the four platforms, integrated search, and metadata harmonization using a common classification system (the UNESCO Thesaurus), serving as a foundation for further technical alignment.

In other recent work, BioPortal upgraded its platform stack, improved multilingual search, and added API functionality for remote terminology pulls. AgroPortal introduced features such as a URI management service and embedded SPARQL editor. EcoPortal implemented a multilingual UI overhaul and structured documentation, while EarthPortal deepened connections with research infrastructures. BiodivPortal expanded its terminology collection, prototyped version diff views, and began exploring LLM integration for enhanced annotation. The OntoPortal Alliance will promote the sharing of these new features across all repositories in the consortium.

OntoPortal provides a shared documentation platform, accessible via ontoportal.github.io. It offers a modular system allowing each portal to present local variations while maintaining a common documentation structure.

For the data community, OntoPortal offers a proven model for scalable terminology repository development, extensible tooling, and metadata interoperability grounded in real-world deployments. The work contributes a blueprint for how FAIR-aligned infrastructure can be adapted and federated across scientific domains.

Deployment and Applications

OntoPortal powers a growing number of public repositories across disciplines. These include BioPortal for biomedical ontologies, AgroPortal for agri-food data, EcoPortal for ecological sciences, EarthPortal for Earth science, and BiodivPortal for biodiversity research. Additional deployments, such as MatPortal and Industry-Portal, demonstrate the software's adaptability. Newer initiatives include OntoPortal-Astro for astronomy, CHPortal for cultural heritage, and LovPortal for Semantic Web vocabularies. Each portal deployment group tailors the software to its community's needs while contributing to shared development across the OntoPortal Alliance.

Key Features

OntoPortal enables users to publish and explore controlled terminologies with rich metadata, track terminology evolution, and reuse terminological content across applications. Mappings of terms across vocabularies help users to identify conceptual relationships, while an Annotator service supports terminology-based text processing. The system provides usage statistics, API access, and community interaction features such as user notes and change proposals. Federation capabilities allow OntoPortal instances to interoperate, sharing artifacts and metadata while retaining local autonomy.

The OntoPortal codebase is open-source and maintained on GitHub. Researchers, developers, and infrastructure projects are invited to deploy new instances, contribute enhancements, or join the OntoPortal Alliance. Our goal is to make ontology repository development as accessible and collaborative as the web itself.

Conclusion

OntoPortal provides an adaptable platform for managing controlled terminologies across disciplines, supported by a growing community of engaged investigators who are tailoring the system for use in a wide range of research domains. The OntoPortal Alliance is reducing the technical and organizational barriers to building sustainable, FAIR-aligned terminology infrastructure. By supporting a federated model with shared tooling and community governance, the OntoPortal platform enables interoperability across domains and the broader Semantic Web ecosystem. As more communities adopt OntoPortal, its role as shared infrastructure continues to expand—lowering barriers to terminology reuse, accelerating adoption of FAIR principles across diverse domains, and offering a practical foundation for semantic services that support open science, linked data, and AI integration.

Primary author: GONÇALVES, Rafael S. (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA)

Co-authors: VENDETTI, Jennifer L. (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA); SKRENCHUK, Alex (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA); DORF, Michael (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA); BOUAZ-ZOUNI, Syphax (Laboratory of Computer Science, Robotics and Microelectronics of Montpellier (LIRMM), Montpellier, France); JONQUET, Clement (French National Research Institute for Agriculture, Food and Environment (INRAE), Paris, France); MUSEN, Mark A. (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA); MEMBERS, OntoPortal Alliance

Presenter: MUSEN, Mark A. (Stanford Center for Biomedical Informatics Research (BMIR), School of Medicine, Stanford University, Stanford, USA)

Session Classification: Presentations Session 5: Rigorous, responsible and reproducible science in the era of FAIR data and AI / Infrastructures to Support Data-Intensive Research

Track Classification: SciDataCon2025 Specific Themes: Open research through Interconnected, Interoperable, and Interdisciplinary Data