

Ontology-based Tools to Enhance the Curation Workflow

Patricia L. Whetzel, Nigam H. Shah, Natasha F. Noy, Clement Jonquet, Cherie Youn,
Paul Alexander, Michael Dorf, and Mark A. Musen

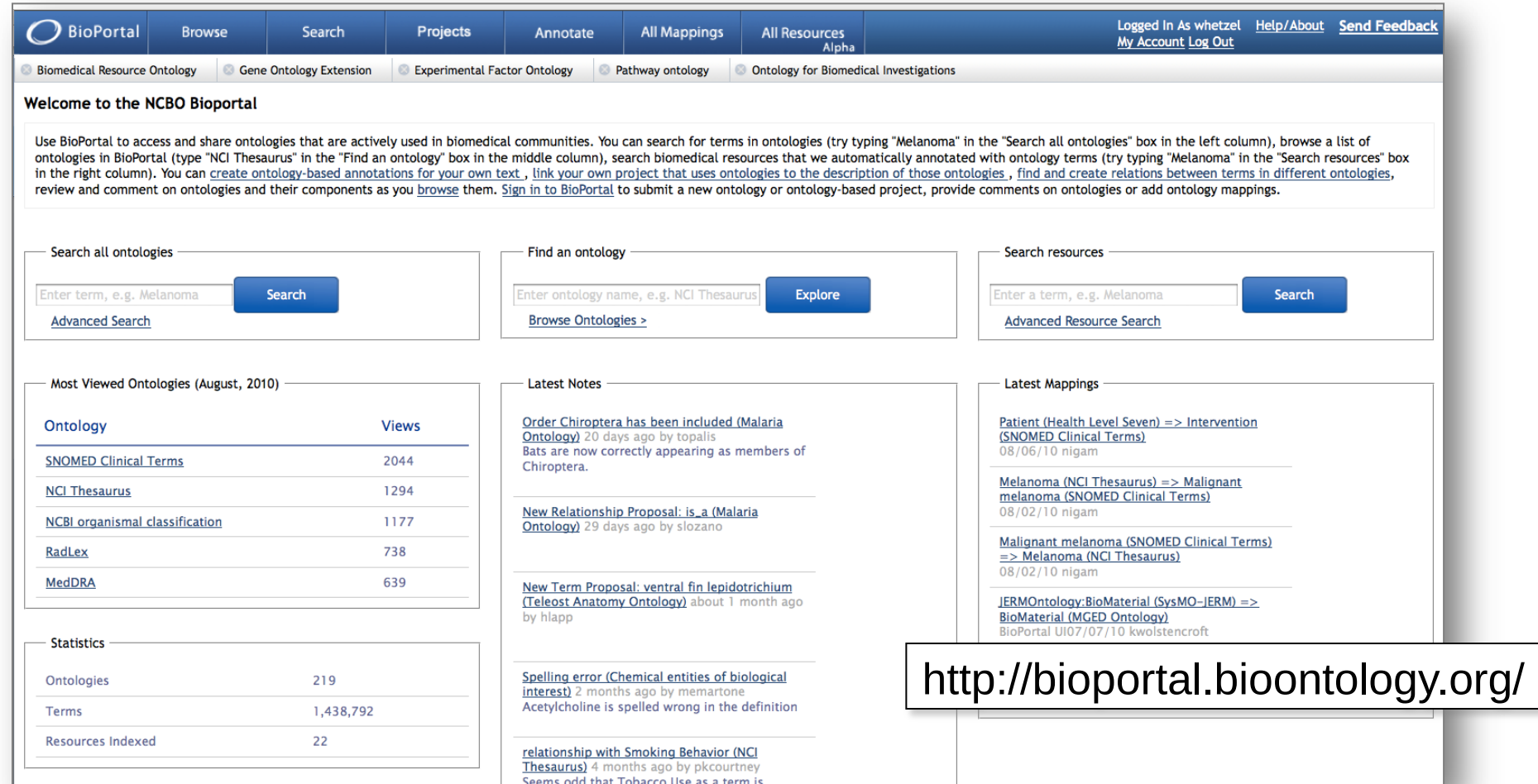
Stanford University, Stanford, CA, USA

Introduction

In order to effectively search, retrieve, and analyze data oftentimes it is curated and tagged with ontology terms. However, the amount of effort to curate the existing set of data resources is beyond the limits of purely manual curation. We present three ontology-based tools developed by the National Center for Biomedical Ontology to enhance the curation workflow: Ontology Widgets, Notes, and the Annotator. The Ontology Widgets provide a mechanism to use ontologies in Web-based forms without the need to locally parse and store the ontology. The Ontology Widgets provide a variety of functionality including term autocompletion and ontology visualization. The Ontology Widgets are implemented for all BioPortal ontologies, including those from the OBO Foundry and Unified Medical Language System. The Notes feature of BioPortal allows structured term proposals to be submitted in order to request the addition or modification of a term in an ontology. The term proposals can be added directly via the BioPortal Web interface or programmatically via the Notes Web service. Notification of new Notes and replies are both RSS- and Email-enabled. Once the term curation process is complete, the OWL class or OBO stanza can be generated via the Notes Web service. Finally, the Annotator can be used to automatically process textual metadata to identify ontology terms found within the text. The Annotator can be accessed programmatically via the Annotator Web service and can be used with all BioPortal ontologies. In summary, the Ontology Widgets, Notes, and Annotator provide mechanisms to enhance curation by helping collect annotated data upon data submission, by facilitating ontology term curation, and by tagging unstructured textual data with ontology terms.

BioPortal Ontology Library

- Open ontology repository for publishing biomedical ontologies
- Features
 - Search across and within all ontologies
 - Browse ontologies
 - Provide feedback – Notes and Reviews
- Statistics
 - Total number of ontologies: 219
 - Number of classes/types: 1,438,792
- Supports many ontology formats: OWL, OBO format, Rich Release Format, Protégé frames



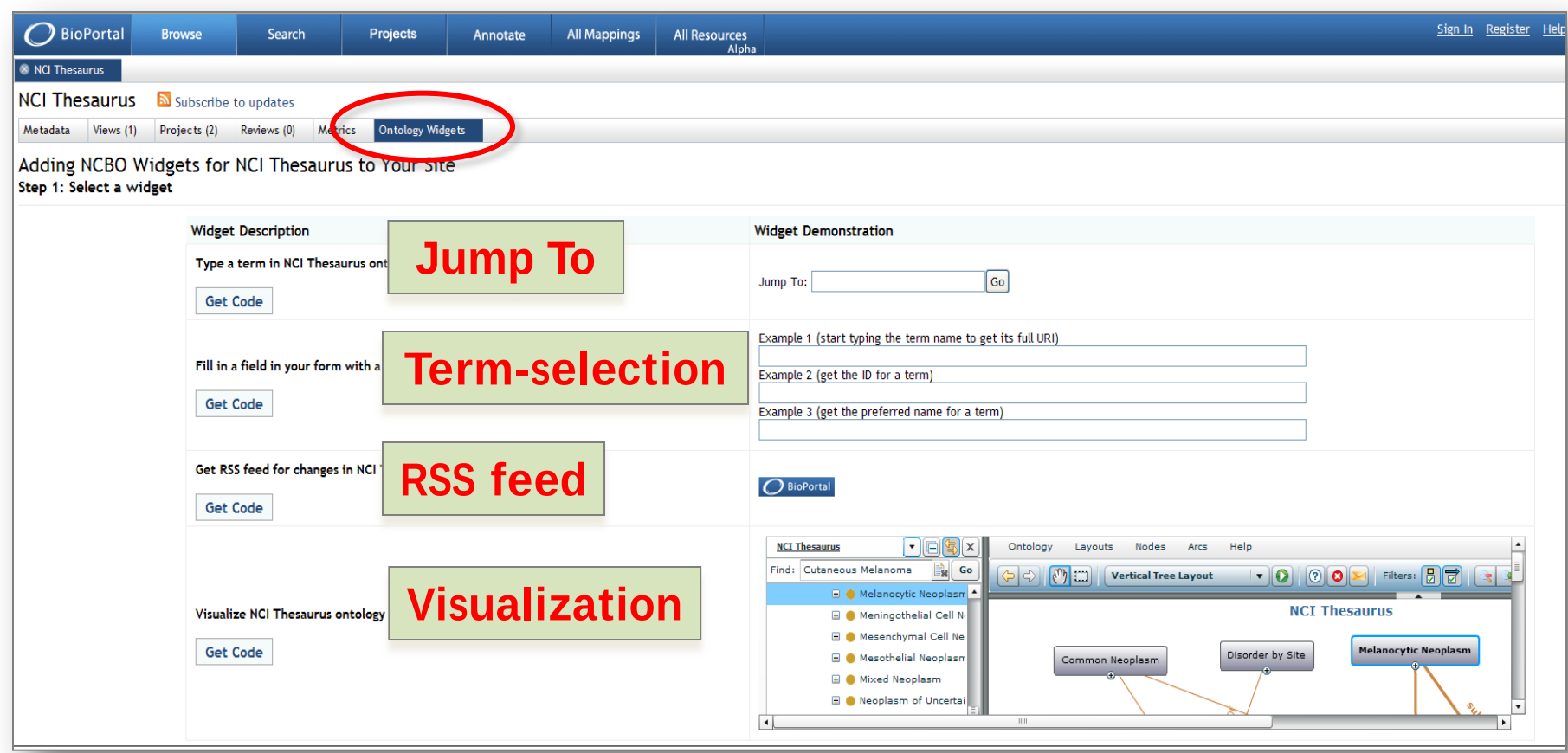
<http://bioportal.bioontology.org/>

Enhancing the Curation Workflow

- Data Submission
 - Ontology Widgets
 - Ontology Web Services
- Ontology Enrichment
 - BioPortal Notes
- Data Annotation
 - NCBO Annotator

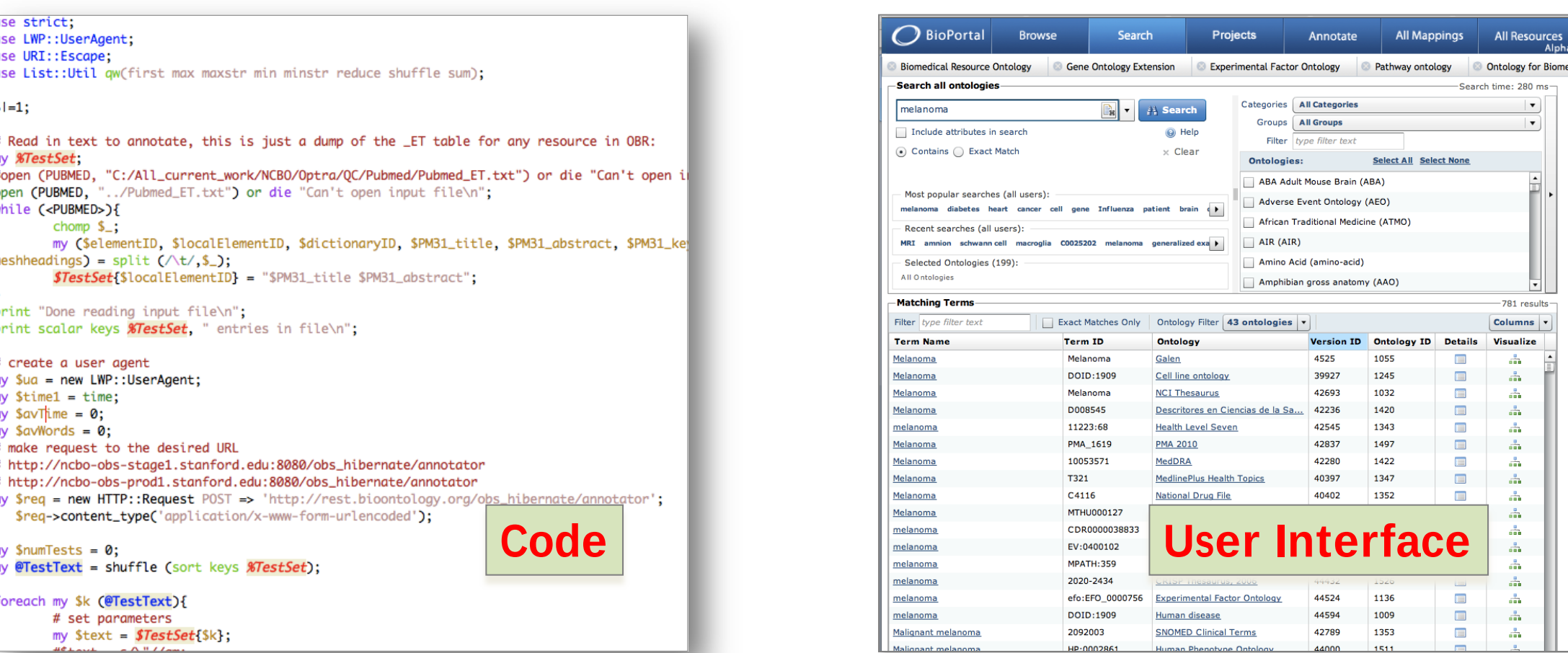
Ontology Widgets

- Ontology widgets provide ready-to-use code to access your ontology of interest from your Web site
- For more information, see:
http://www.bioontology.org/wiki/index.php/Ontology_Widgets



Web Service Access

- Prefix for NCBO Web Services:
<http://rest.bioontology.org/{application}>
- List of Web Services
http://bioontology.org/wiki/index.php/NCBO_REST_services
- Access via Code (Perl, Java, Flash, Ruby, etc.), User Interface, Browser, or Widget



Acknowledgements

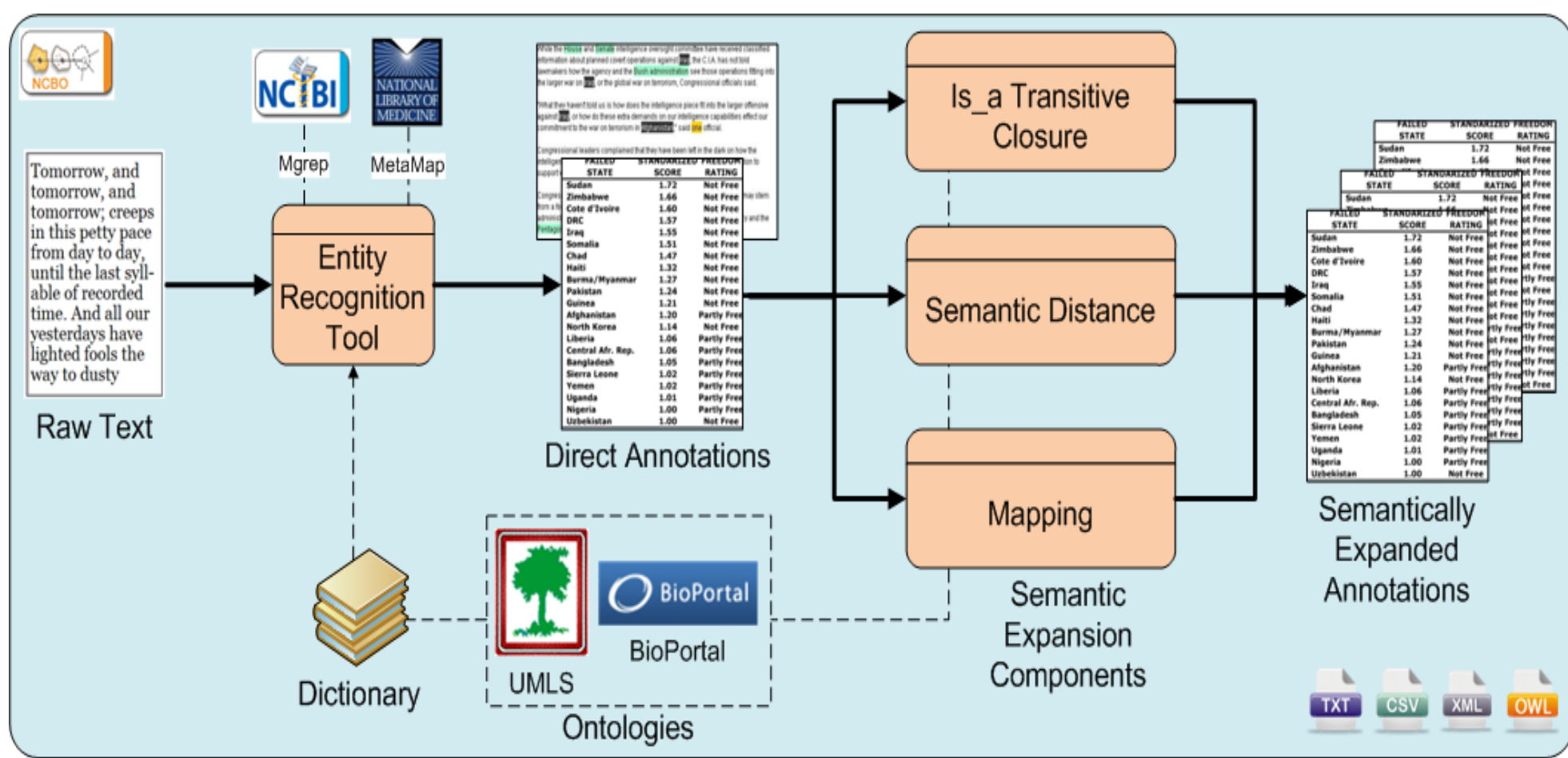
BioPortal is developed by the National Center for Biomedical Ontology (NCBO), a National Center for Biomedical Computing under the NIH Roadmap. BioPortal is developed in conjunction with partners at the University of Victoria and the Mayo Clinic. For more information on NCBO see <http://www.bioontology.org>.

BioPortal Web Services

- Prefix for BioPortal Web Services
<http://rest.bioontology.org/bioportal/>
- Ontology Web services
 - List metadata for all ontologies
 - List all versions of an ontology
 - Download ontologies
 - List all ontology categories
 - List all ontology groups
- Search Web service
 - Search for terms from all BioPortal Ontologies
 - Search for term within a subtree of an ontology
- Term Web services
 - Get term by id
 - Get all root terms for an ontology
 - Get all terms from an ontology
- Hierarchy Web services
 - Get parents/children of a term
 - Get path to roots/leaves of a term
 - Get siblings of a term
- RDF Generation Web service
 - Generate a N3 RDF file
- Notes Web service
 - Get comments for a term
 - Get all comments for an ontology
- Mapping Web service
 - Get mappings for a term
 - Get all mappings for an ontology

Annotator Web Service

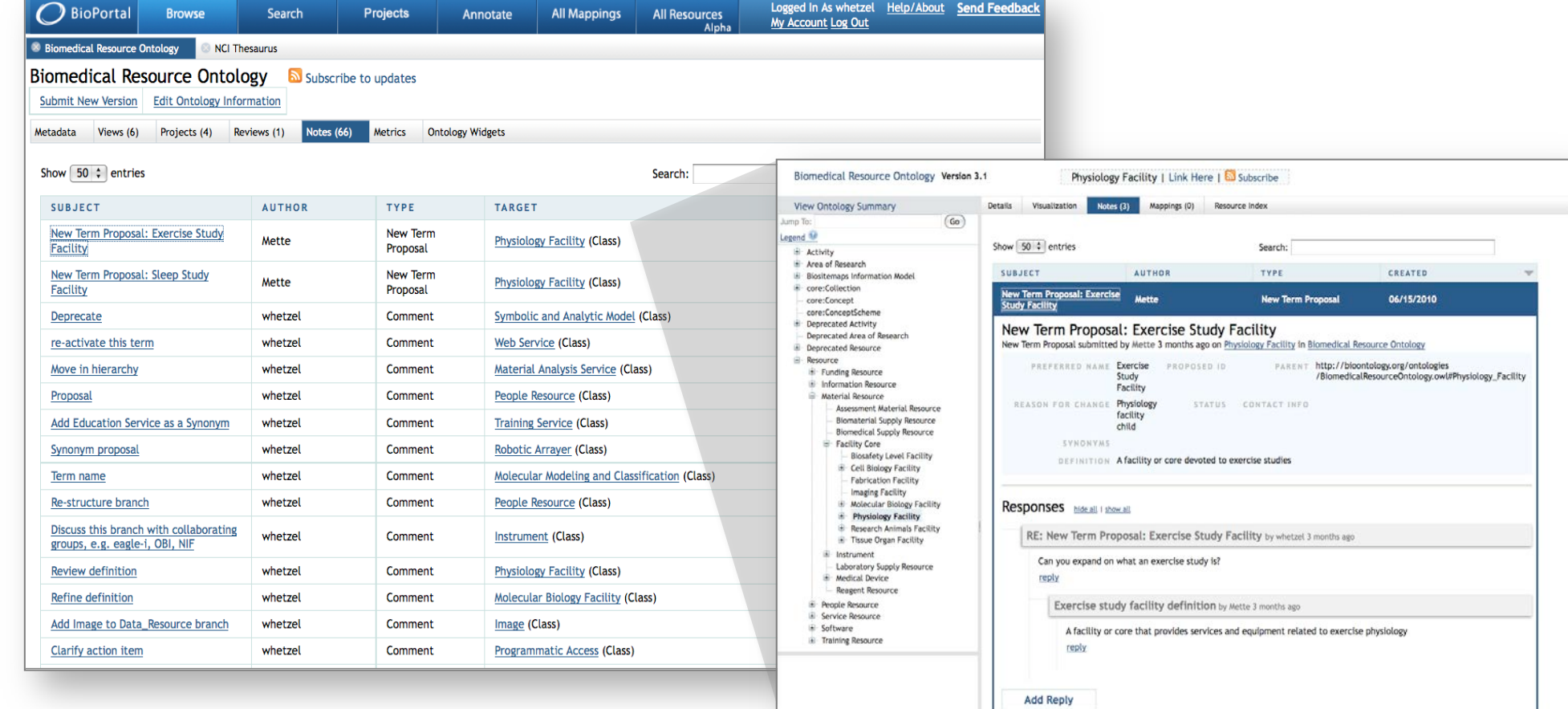
- Prefix for Annotator Web Service
<http://rest.bioontology.org/obs/annotator/>
- Description
 - The NCBO Annotator annotates or “tags” free-text data with terms from BioPortal and UMLS ontologies. The current entity recognizer used by the Annotator is Mgrep, developed by the National Center for Integrative Biomedical Informatics. Plans are underway to include an additional entity recognizer, MetaMap from the National Library of Medicine. The annotation is performed in two steps; direct annotations via matching the input text with the preferred name or synonyms of ontology terms and expanded annotations taking into account the ontology hierarchy and mappings.
- Documentation
 - For more information and sample clients, see:
http://www.bioontology.org/wiki/index.php/Annotator_Web_service



Ontology Development Lifecycle

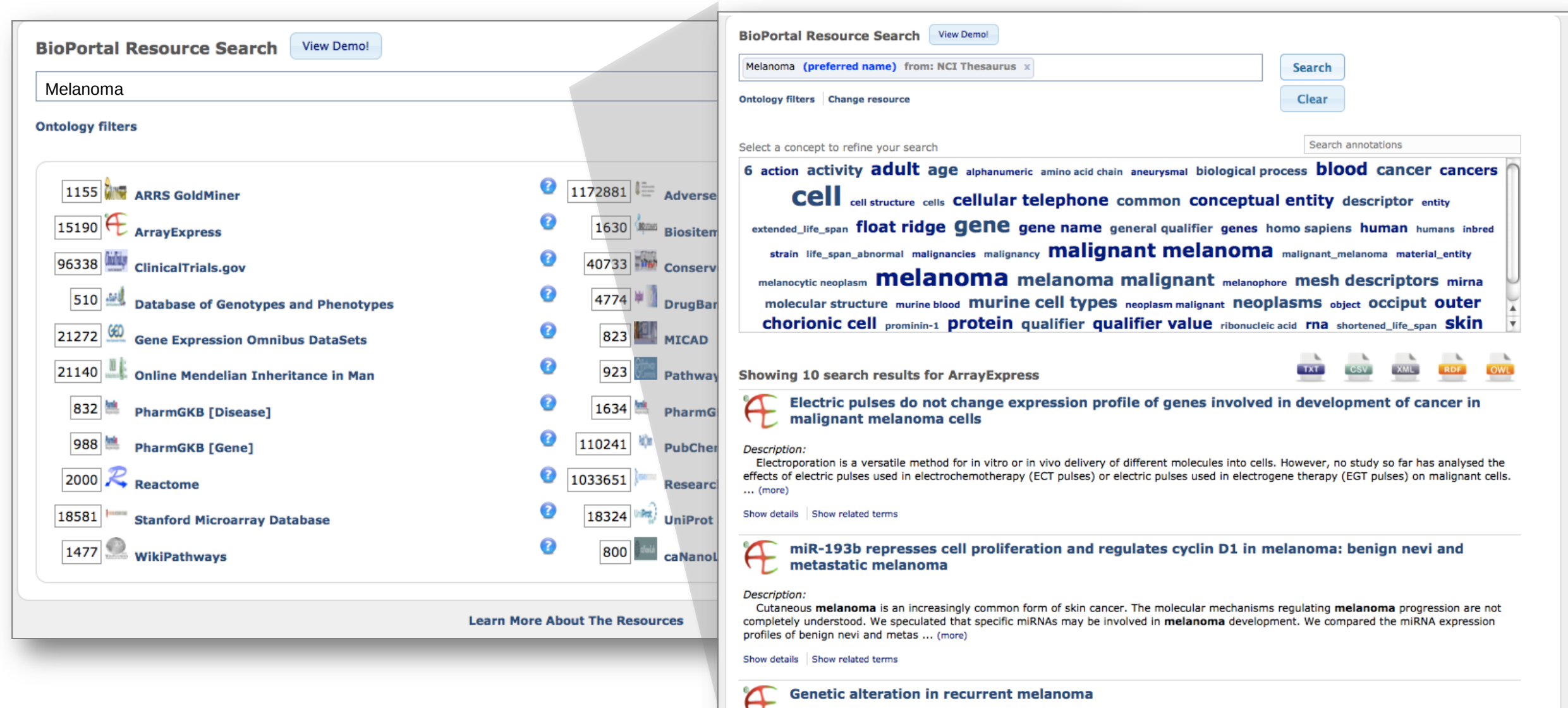
- Involves both Subject Matter Experts (SMEs) and Ontology Developers
- Consists of:
 - Discussion between SMEs and Ontology Developers
 - Ontology Editing
 - Publishing Ontology to the Web
 - Collect Feedback on Published Ontology

BioPortal Notes



Resource Index Web Service

- Prefix for Resource Index Web Service
http://rest.bioontology.org/resource_index/
- Description
 - Provides an ontology-based index of publicly available biomedical resources.
 - Currently indexed resources include: Array Express, ARRS Goldminer, Biositemaps, caNanoLab, ClinicalTrials.gov, Conserved Domain Database, DrugBank, Gene Expression Omnibus, Molecular Imaging and Contrast Agent Database, NextBio, OMIM, Pathway Commons, PharmGKB, PubChem, Reactome, Research Crossroads, Stanford Microarray Database, UniProt, and WikiPathways
- For more information see:
http://www.bioontology.org/wiki/index.php/Resource_Index



Contact

BioPortal Support: support@bioontology.org
Protégé/WebProtege: protege-discussion@lists.stanford.edu