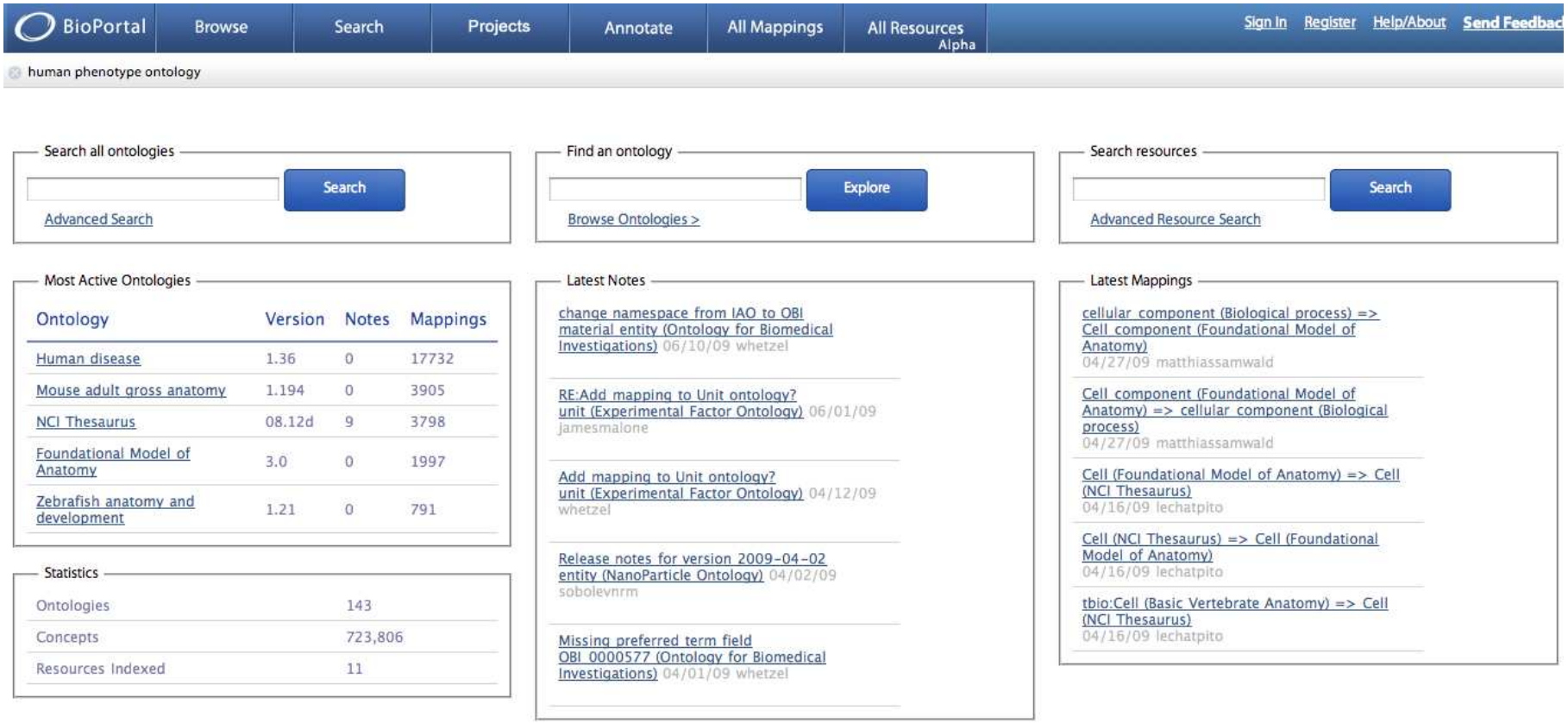
Mapping Repository in BioPortal

Natasha Noy
Stanford University

BioPortal: A Community-Based Ontology Repository



http://bioportal.bioontology.org

BioPortal

- Web accessible repository of ontologies for the biomedical community:
  - http://bioportal.bioontology.org
- Online support for ontology
  - Peer review
  - Notes (comments and discussion)
  - Versioning
  - Mapping
  - Search
  - Resources

The BioPortal Technology

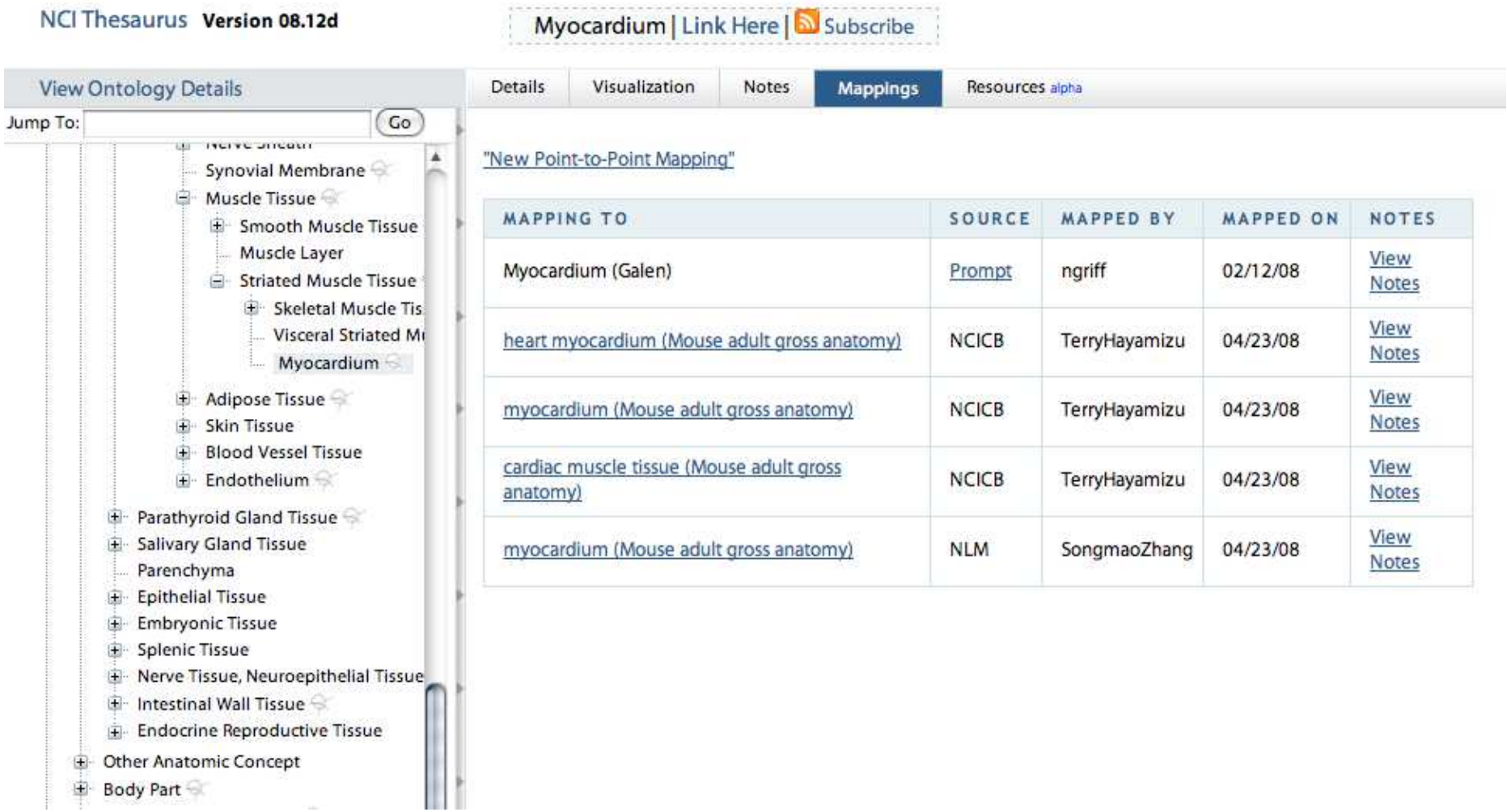
- BioPortal is a library of more than 300 biomedical ontologies
- All BioPortal data is accessible through REST services
  - BioPortal user interface accesses the repository through REST services as well
- The BioPortal technology is domain-independent
  - There are installations of BioPortal for libraries in other domains
  - BioPortal code is open-source

Why Would a Mapping Repository Help?

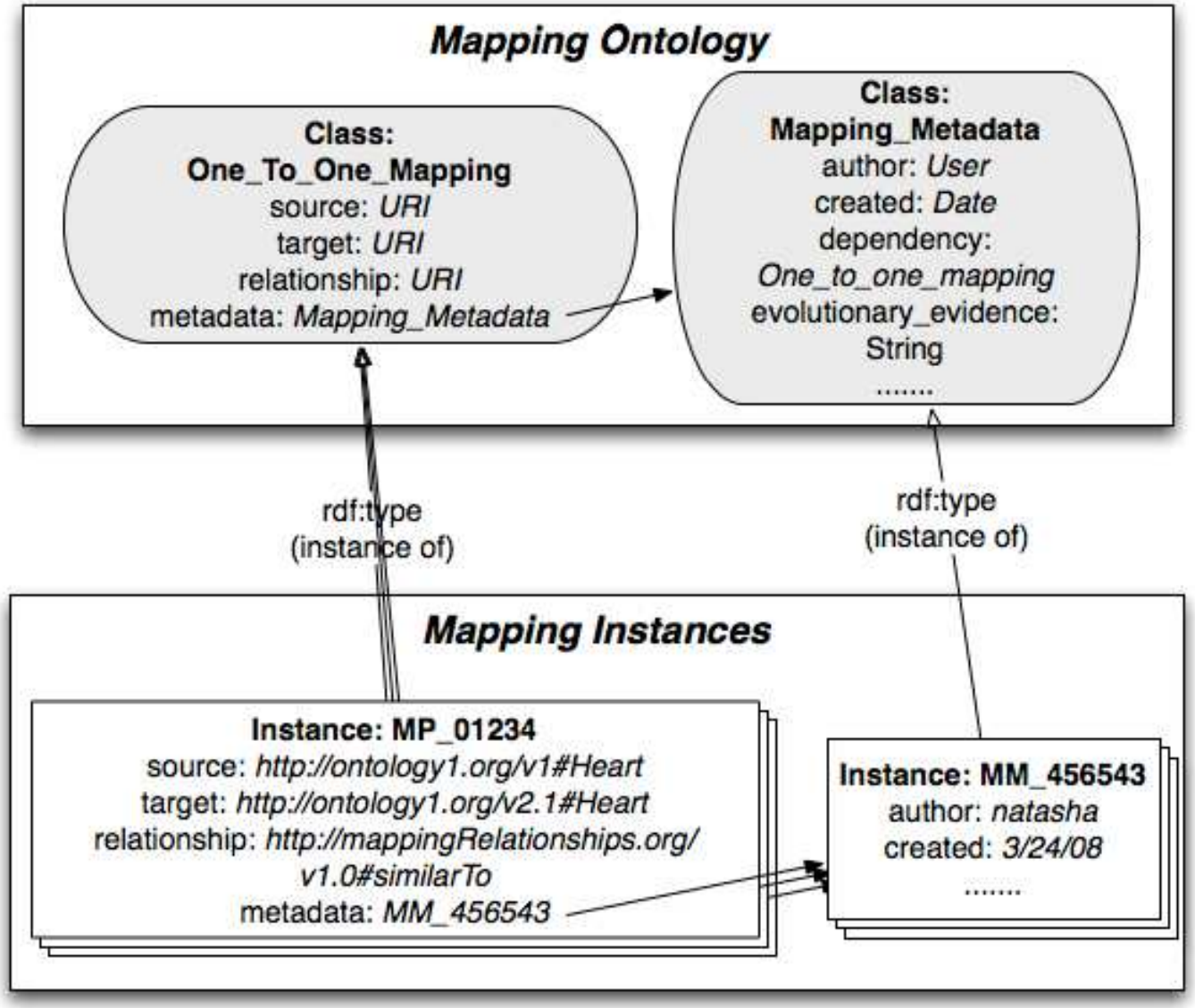
- Source of data for automatic algorithms
  - machine learning
  - algorithms that need a priori alignment
- Use for annotating and browsing resources through ontology elements
- Use for finding "important" ontologies:
  - If everyone maps to NCI Thesaurus, it must be important
- Accessible through web services
  - can be used in other applications

Mappings in BioPortal

- Mappings in BioPortal are concept-to-concept mappings
- Mappings are created by users or uploaded in bulk
- Bulk uploads are usually the results of automatic or semi-automatic mapping
- There is detailed metadata for provenance of mappings
- ~2M mappings in BioPortal now


NCI Thesaurus Version 08.12d
Myocardium | Link Here | Subscribe
View Ontology Details
Jump To:
Go
Synovial Membrane
Muscle Tissue
Smooth Muscle Tissue
Muscle Layer
Striated Muscle Tissue
Skeletal Muscle Tis
Visceral Striated M
Myocardium
Adipose Tissue
Skin Tissue
Blood Vessel Tissue
Endothelium
Parathyroid Gland Tissue
Salivary Gland Tissue
Parenchyma
Epithelial Tissue
Embryonic Tissue
Splenic Tissue
Nerve Tissue, Neuroepithelial Tissue
Intestinal Wall Tissue
Endocrine Reproductive Tissue
Other Anatomic Concept
Body Part
Details
Visualization
Notes
Mappings
Resources alpha
"New Point-to-Point Mapping"
MAPPING TO
SOURCE
MAPPED BY
MAPPED ON
NOTES
Myocardium (Galen)
Prompt
ngriff
02/12/08
View Notes
heart myocardium (Mouse adult gross anatomy)
NCICB
TerryHayamizu
04/23/08
View Notes
myocardium (Mouse adult gross anatomy)
NCICB
TerryHayamizu
04/23/08
View Notes
cardiac muscle tissue (Mouse adult gross anatomy)
NCICB
TerryHayamizu
04/23/08
View Notes
myocardium (Mouse adult gross anatomy)
NLM
SongmaoZhang
04/23/08
View Notes

Representing mappings



Mapping Metadata

- mapping relationship
- provenance (who created the mapping and when)
- comments, which may include transformation scripts, etc.
- discussion and comments
- application context
- mapping dependency
- algorithm used to create the mapping (configuration, parameters, etc.)
- external references

Types of mappings in BioPortal

- Lexical match on preferred names or preferred names and synonyms (skos:closeMatch)
- For ontologies from UMLS, a CUI-based mapping
  - created if two terms have the same CUI
  - skos:closeMatch
- Terms in different ontologies with the same URI (skos:exactMatch)

Issues

- Mappings and ontology versions
  - Re-generate all automatically generated mappings as new versions become available
  - Delete mappings for previous versions, unless they have discussions associated with them
  - Make mappings accessible through all versions
- Scalability
  - Majority of mappings have the same provenance data
    - Reuse the object that represents the provenance metadata

Current work

- Re-generating all the mappings
- Adding mappings to the SPARQL endpoint