How to Organize the World of Ontologies

Barry Smith

http://ontology.buffalo.edu/smith

Why build ontologies?

To solve the data silo problem – there are too many ways to create terminologies and databases

We need to *constrain* terminologies and databases so that they converge

Make them conform to a single evolving consistent set of ontologies covering the whole of reality

Make all these ontologies conform to a common set of tested guidelines

NCOR

National Center for Ontological Research, Buffalo

Core ontologies and associated development guidelines:

Basic Formal Ontology (BFO) ,2002-

Relation Ontology (RO), 2004-

Ontology for Biomedical Investigations (OBI), 2005-

Information Artifact Ontology (IAO), 2008

NCOR goals

Formulate and test guidelines

- for building ontologies
- for linking ontologies
- for evaluating ontologies
- for applying ontologies

Establish and disseminate best practices

OBO Foundry

Ontology development guidelines being tested in a large community of users of ontologies in addressing the retrieval and integration biomedical data

Model now being followed also e.g. in NIH Neuroscience Information Framework Foundry, in MIBBI (Minimal Information about a Biological and Biomedical Investigation) Foundry

A success story in information integration

OBO Foundry network of interoperable ontology modules (http://obofoundry.org)

All modules configured as extensions of BFO as common top-level semantic layer simple enough to be used by biologists who are not IT experts

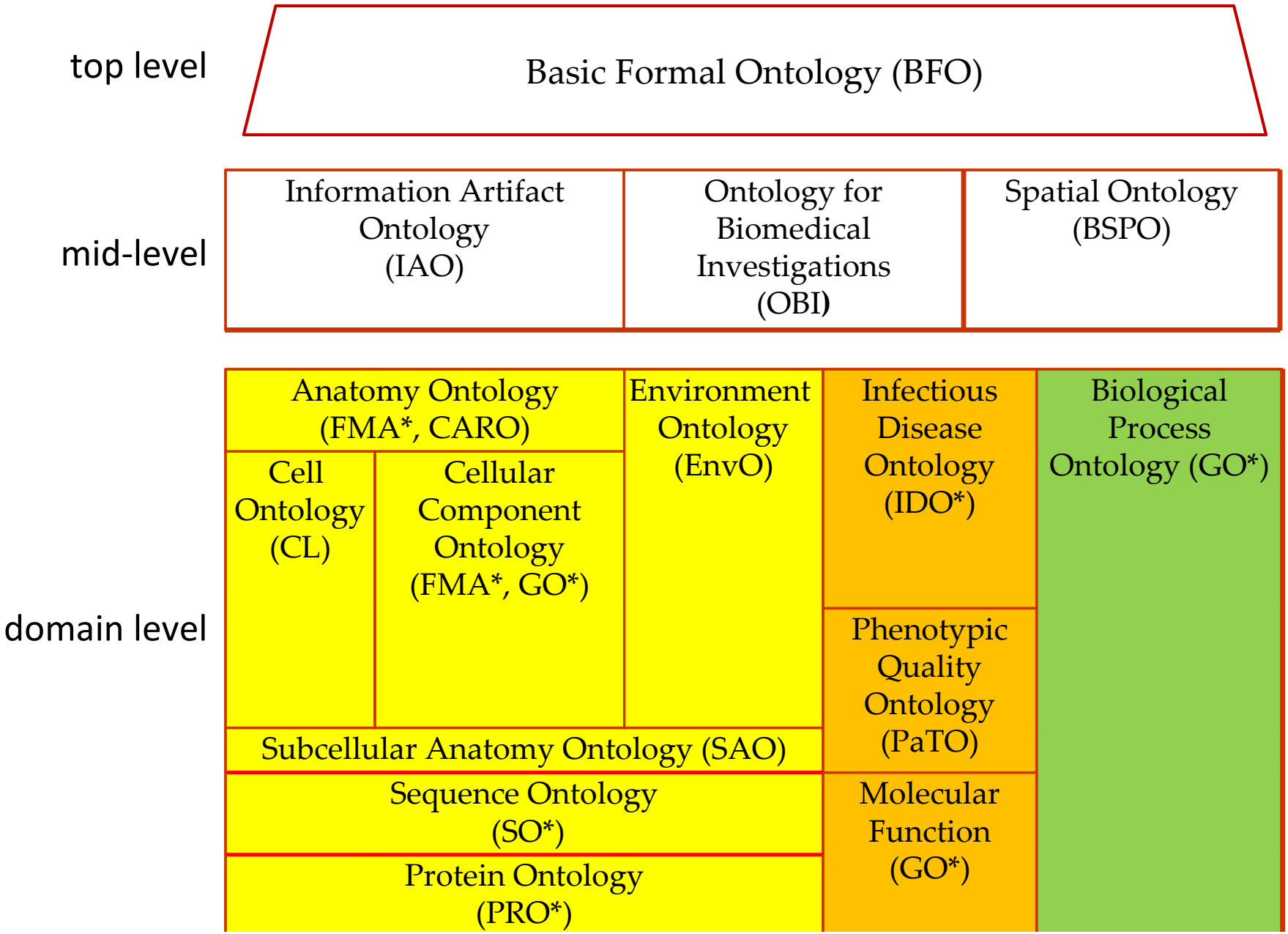
All modules subjected to joint evolution and peer review

Used by 1000s of researchers to promote semantic interoperability of experimental data in scores of high-throughput domains of biology and medicine

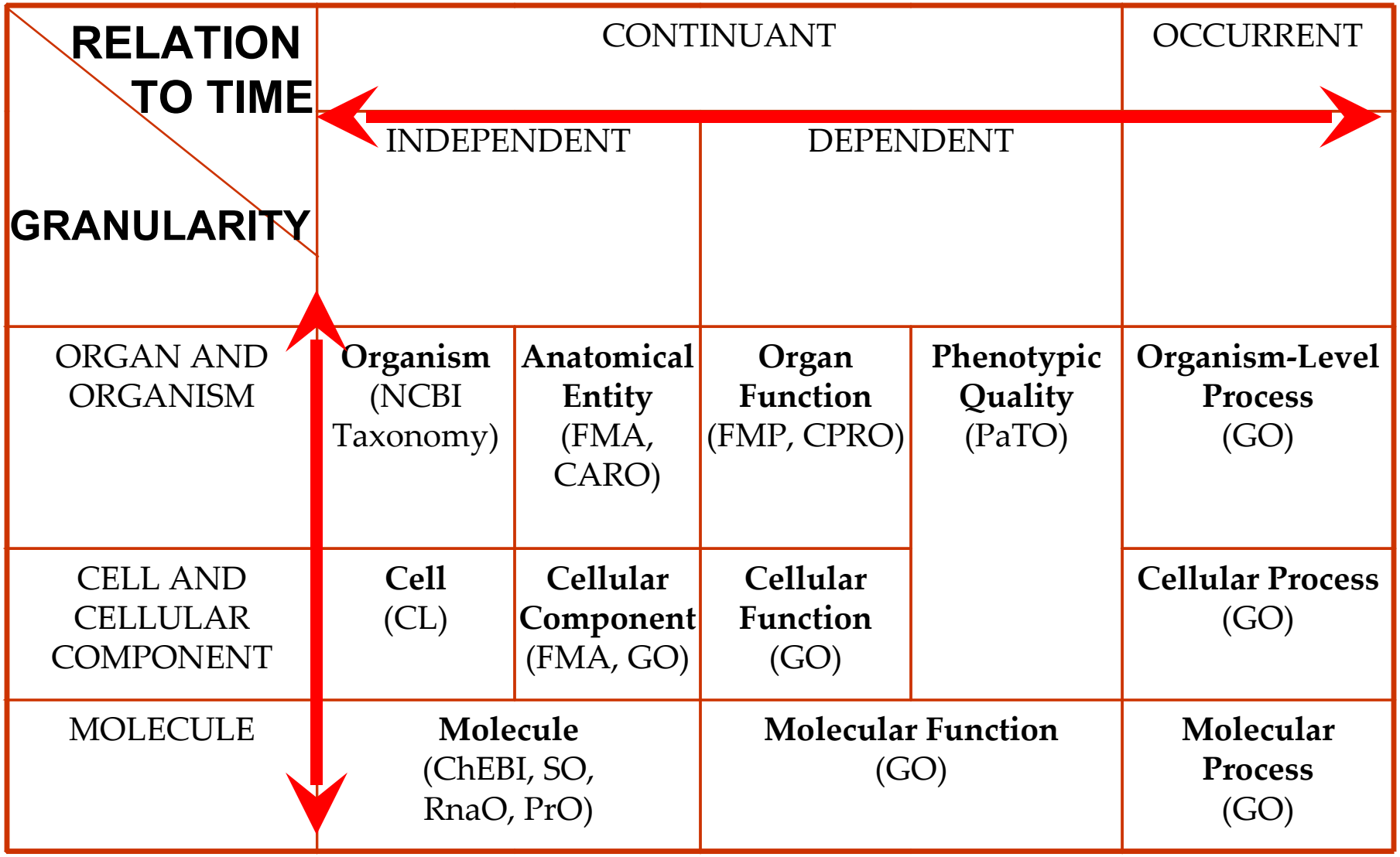
Ontologists are abandoning local ontologies to support common resources

Unifying goal: integration of data

- within and across domains
- across different species
- across levels of granularity (organ, organism, cell, molecule)
- across different perspectives (physical, biological, clinical)



OBO Foundry Modular Organization



obofoundry.org

Principal BFO Types

CONTINUANT (endures through time ≈ UCore "Entity")			OCCURRENT (occurs in time ≈ UCore "Event")	
INDEPENDENT	DEPENDENT	SPATIAL	PROCESS	TEMPORAL
Object: Person, Rock, Vehicle	Attribute: Quality, Role, Capability	Spatial Region	Speaking, Walking, Flying	Temporal Interval, Spatiotemporal Region

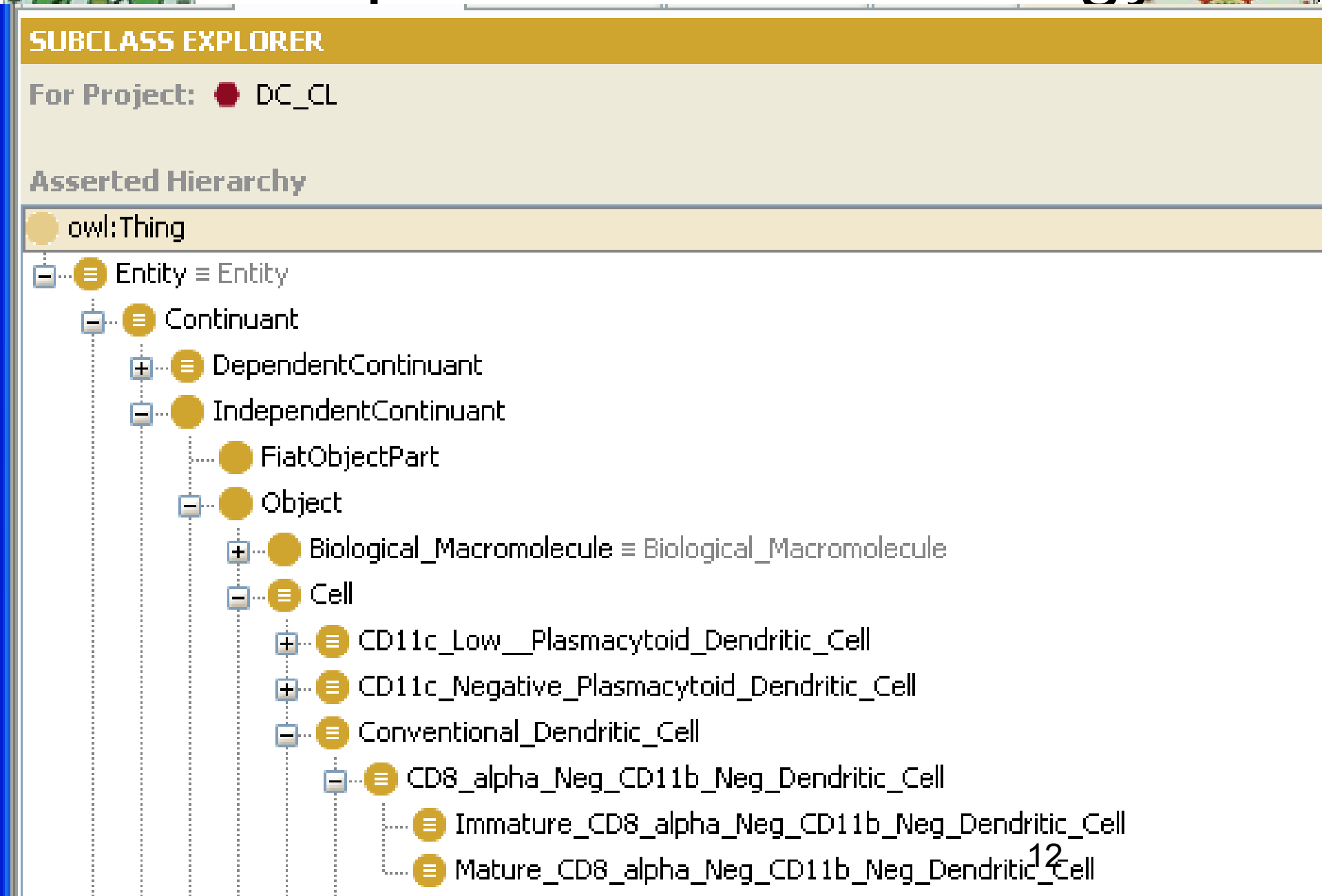
Two Examples

OBI: Ontology for Biomedical Investigations

IDO: Infectious Disease Ontology

CL: The Cell Ontology

Example: The Cell Ontology



OBI Collaborating Communities

Environmental Genomics MGED RSBI Group
Genomic Standards Consortium (GSC)
HUPO Proteomics Standards Initiative (PSI)
Immunology Database and Analysis Portal
Immune Epitope Database and Analysis Resource (IEDB)
International Society for Analytical Cytology
Metabolomics Standards Initiative (MSI),
Neurogenetics, Biomedical Informatics Research Network (BIRN)
Nutrigenomics MGED RSBI Group
Toxicogenomics MGED RSBI Group
Transcriptomics MGED Ontology Group

IDO (Infectious Disease Ontology) Consortium

MITRE, Mount Sinai, UTSouthwestern – Influenza

IMBB/VectorBase – Vector borne diseases (*A. gambiae, A. aegypti, I. scapularis, C. pipiens, P. humanus*)

Colorado State University – Dengue Fever

Duke University – Tuberculosis, *Staph. aureus*

Cleveland Clinic – Infective Endocarditis

University of Michigan – Brucilosis

University of Michigan – Vaccine Ontology

Three criteria of a successful standard

1. intelligibility to users, consistent use of terms like ‘term’, ‘class’, ‘entity’, ‘object’ …)
2. track record of lessons learned (GO has 10 years of hard user testing)
3. lots of existing users (ontologies are like telephone networks)