

COB: A Core Ontology for Biology



Ontology Summit 2023

<https://ontologyforum.org/index.php/OntologySummit2023>

Chris Mungall

Lawrence Berkeley National Laboratory



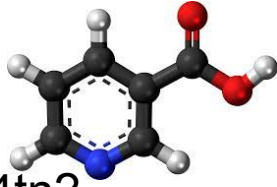
<https://genomic.social/@cmungall>



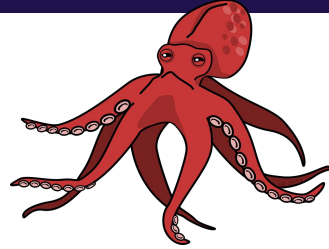
@chrismungall

We have many identifiable things and categories

Drugs 10k
Chemicals 4tn?



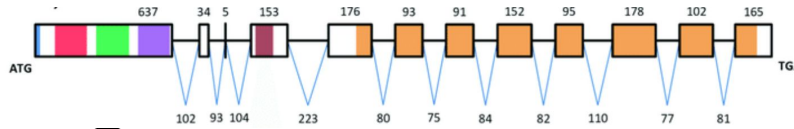
Species
~9 million



Cells
10,000s+
types
per species)



Genes 20k per species
Metagenome dbs: **>65bn** genes



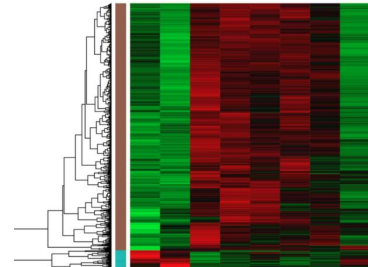
Genetic
variants
3m in human
alone



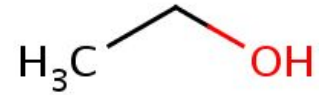
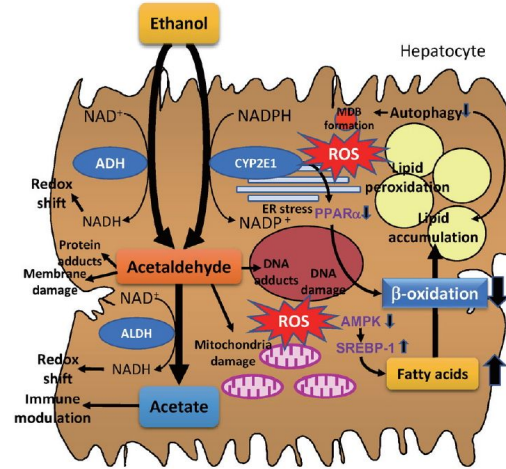
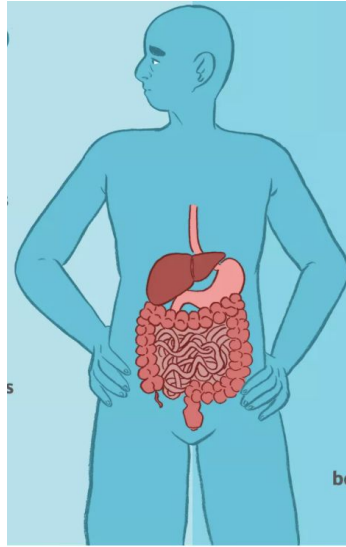
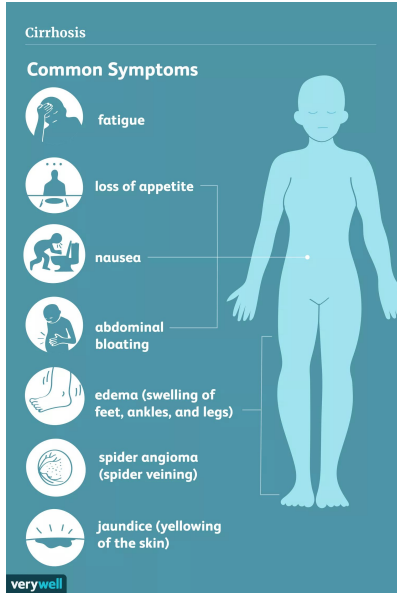
Diseases and
Phenotypes
10-50k/species



Experiments
Raw data
?? exabytes



The things are interconnected across scales



Different perspectives on the same things

static

disease state

static

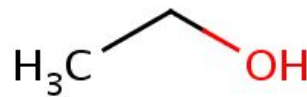
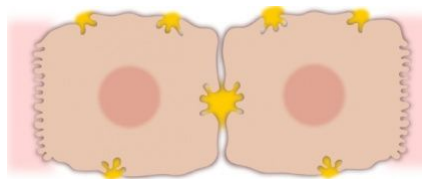
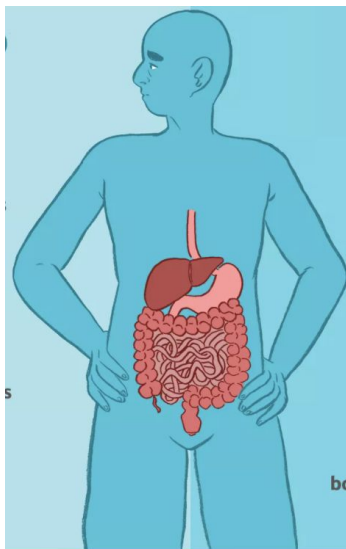
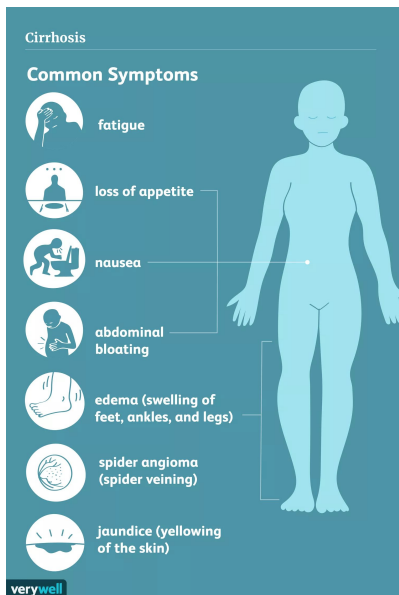
anatomy

static

cell structures

static

chem structures



Different perspectives on the same things

dynamic

pathophysiology

dynamic

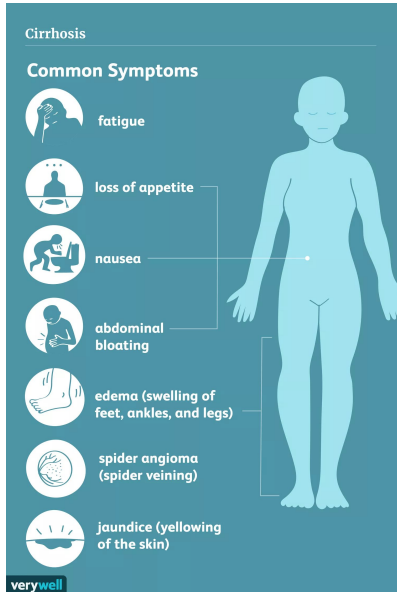
physiology

dynamic

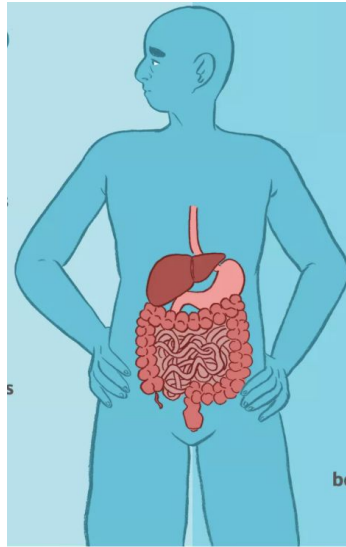
pathways

dynamic

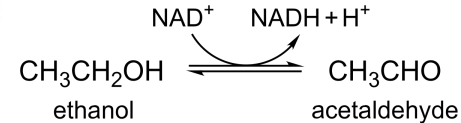
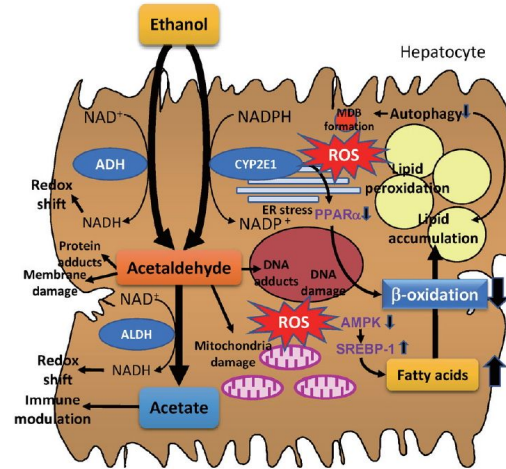
reactions



↑bilirubin → jaundice

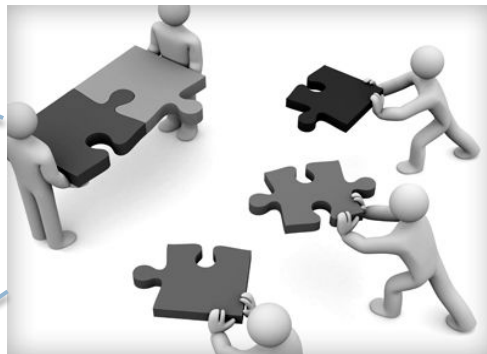


Conjugation of bilirubin occurs-in liver



OBO is for organizing the things

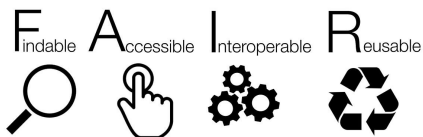
1. Well-integrated
Modular ontologies



2. **Technical and sociotechnological framework** for cooperation

4. Allow us to categorize, organize, and **integrate** all of the things

3. **Tools, best practices and infrastructure** for forging new ontologies



 @obofoundry

<http://obofoundry.org>



OBO registry

OBO Library: find, use, and contribute to community ontologies

Download table as: [[YAML](#) | [JSON-LD](#) | [RDF/Turtle](#)]

Search Table

Ontology Domains:



Group By Domain



















































Hide Inactive



Hide Obsolete

anatomy and development

Anatomy and development

ID ^	Title ^	Description	Quick Access	Re-Use ^	Social
aism	Ontology for the Anatomy of the Insect SkeletoMuscular system (AISM)	The AISM contains terms used in insect biodiversity research for describing structures of the exoskeleton and the skeletomuscular system. It...	     		 9
amphx	The Amphioxus Development and Anatomy Ontology	An ontology for the development and anatomy of Amphioxus (Branchiostoma lanceolatum).	     		 3
bspa	Biological Spatial Ontology	An ontology for representing spatial concepts, anatomical axes, gradients, regions, planes, sides, and surfaces	     		 10
bto	BRENDA tissue / enzyme source	A structured controlled vocabulary for the source of an enzyme comprising tissues, cell lines, cell types and cell cultures.	     		 4
caro	Common Anatomy Reference Ontology	An upper level ontology to facilitate interoperability between existing anatomy ontologies for different species	     		 3
cl	Cell Ontology	The Cell Ontology is a structured controlled	     		 100

Ontology Domains:



Grc



agriculture

anatomy and development

biological systems

chemistry and biochemistry

diet, metabolomics, and nutrition

environment

health

information

information technology

investigations

microbiology

organisms

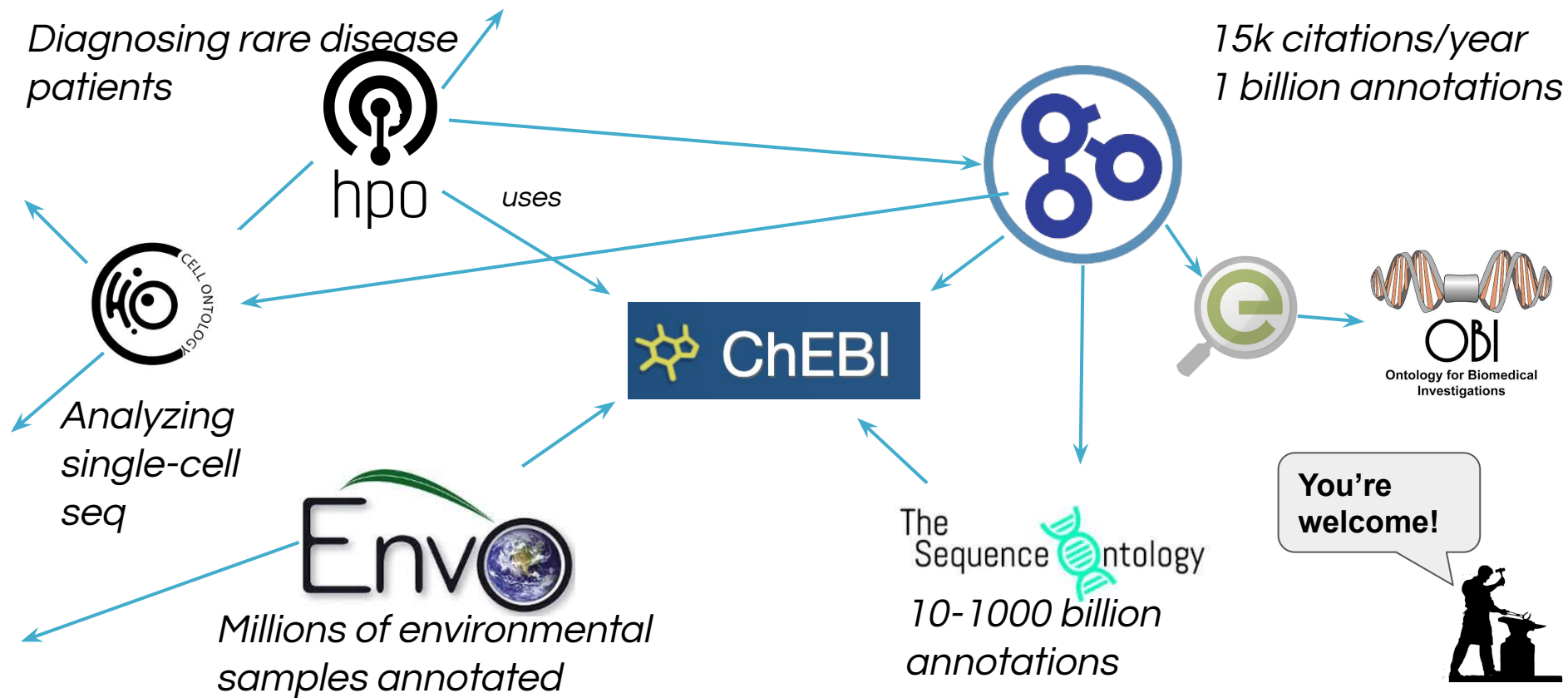
phenotype

simulation

upper

<https://obofoundry.org>

OBO ontologies *interoperate*



Layers of OBO interoperability

— — —

Shared vocabularies and upper level integration

RO standardises the **relationships** to be used in OBO ontologies.

OMO standardises the **annotation properties** to be used for term and ontology metadata.

COB provides the **upper layer** for biological and biomedical ontologies.

Term-reuse across OBO ontologies.

FA[IR]ness & Openness

Ontologies should be findable, accessible and openly available.



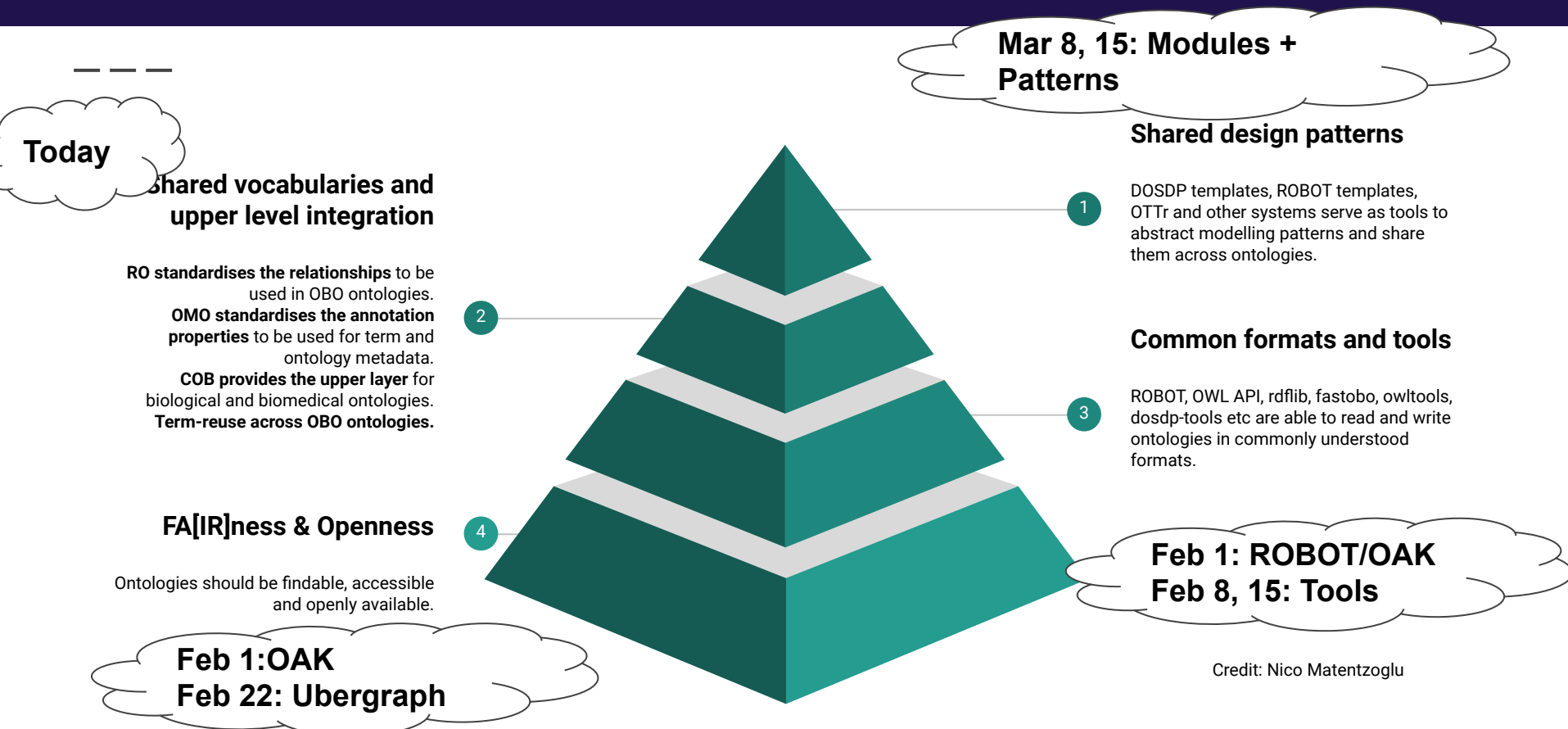
Shared design patterns

DOSDP templates, ROBOT templates, OTTr and other systems serve as tools to abstract modelling patterns and share them across ontologies.

Common formats and tools

ROBOT, OWL API, rdfib, fastobo, owltools, dosdp-tools etc are able to read and write ontologies in commonly understood formats.

Layers of OBO interoperability



Shared upper level integration across OBO

Shared vocabularies and upper level integration

RO standardises the **relationships** to be used in OBO ontologies.

OMO standardises the **annotation properties** to be used for term and ontology metadata.

COB provides the **upper layer** for biological and biomedical ontologies.
Term-reuse across OBO ontologies.

RO: Standardized relationships

- **Relation Ontology**
- How are biological entities related?
- E.g. part-of, develops-from

OMO: Standardized annotation properties

- **OBO Metadata Ontology**
- Metadata about terms and ontologies
- E.g. definition, synonym, creator

COB: common biological upper layer

- **Common Ontology for Biology**
- One layer beneath BFO

BFO:

- **Basic Formal Ontology**
- Layer of abstractions above COB

A brief history, from RO to COB



Relation Ontology (RO), started 2004

Method | [Open Access](#) | [Published: 28 April 2005](#)

Relations in biomedical ontologies

[Barry Smith](#) , [Werner Ceusters](#), [Bert Klagges](#), [Jacob Köhler](#), [Anand Kumar](#), [Jane Lomax](#), [Chris Mungall](#), [Fabian Neuhaus](#), [Alan L Rector](#) & [Cornelius Rosse](#)

[Genome Biology](#) **6**, Article number: R46 (2005) | [Cite this article](#)

53k Accesses | **579** Citations | **18** Altmetric | [Metrics](#)

Abstract

To enhance the treatment of relations in biomedical ontologies we advance a methodology for providing consistent and unambiguous formal definitions of the relational expressions used in such ontologies in a way designed to assist developers and users in avoiding errors in coding and annotation. The resulting Relation Ontology can promote interoperability of ontologies and support new types of automated reasoning about the spatial and temporal dimensions of biological and medical phenomena.

<https://doi.org/10.1186/gb-2005-6-5-r46>

Relation	Transitive	Symmetric	Reflexive
<i>is_a</i>	+	-	+
<i>part_of</i>	+	-	+
<i>located_in</i>	+	-	+
<i>contained_in</i>	-	-	-
<i>adjacent_to</i>	-	-	-
<i>transformation_of</i>	+	-	-
<i>derives_from</i>	+	-	-
<i>preceded_by</i>	+	-	-
<i>has_participant</i>	-	-	-
<i>has_agent</i>	-	-	-

Original RO

OWL formulation of RO

located in — RO:0001025 — http://purl.obolibrary.org/obo/RO_0001025

Annotations Usage

Annotations: located in

Annotations +

[rdfs:label](#) [language: en]
located in

[IAO_0000115](#) [language: en]
a relation between two independent continuants, the target and the location, in which the target is entirely within the location

Characteristic: Description: located in

☐ Functional

☐ Inverse function

☒ Transitive

☐ Symmetric

☐ Asymmetric

Equivalent To +

SubProperty Of +

Inverse Of +

'location of'

OBO Relation Ontology

Search docs

BASICS

Index

Quickstart

History

GUIDE

Introduction

OWL Concepts

Annotation Properties

Object Properties

Property Chains

Identifiers

Shortcut Relations

Examples

Domain and Ranges

DESIGN PATTERNS

Design Patterns Overview

Docs » Design Patterns » Reflexivity

Reflexivity in RO

We first introduce a pattern used in RO for combination with property chains, the general treatment of reflexivity in RO, and how relations in RO are rarely *globally* reflexive.

Defining Property Chains involving Reflexivity

When defining property chains over R and R_2 we typically name the chain and include an axiom:

```
ObjectProperty: {R}_ {R2}  
SubObjectPropertyOf: R o R2
```

we may also want to make this a reflexive property chain:

```
ObjectProperty: R  
SubObjectPropertyOf: {R}_ {R2}
```

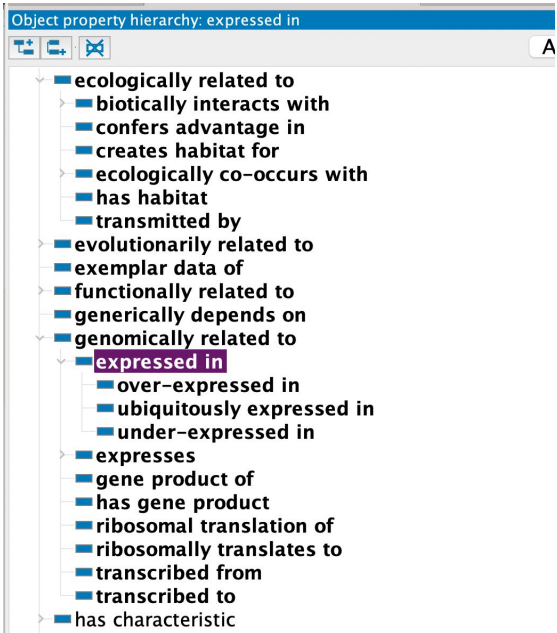
When this OPR is used, the property definition should be based on the reflexive property chain.

<https://oborel.github.io>

Original RO

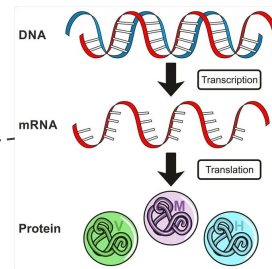
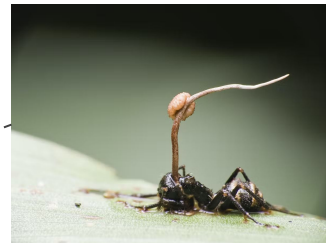
RO
development

RO Core and extensions



RO: **all** relations

- RO-Core: *cross-domain*
 - Part of
 - Located in
 - Occurs in
 - Characteristic of
- Ecological relations
 - Parasite of
 -
- Genomic relations
 - Expressed in
 - ...
- ...

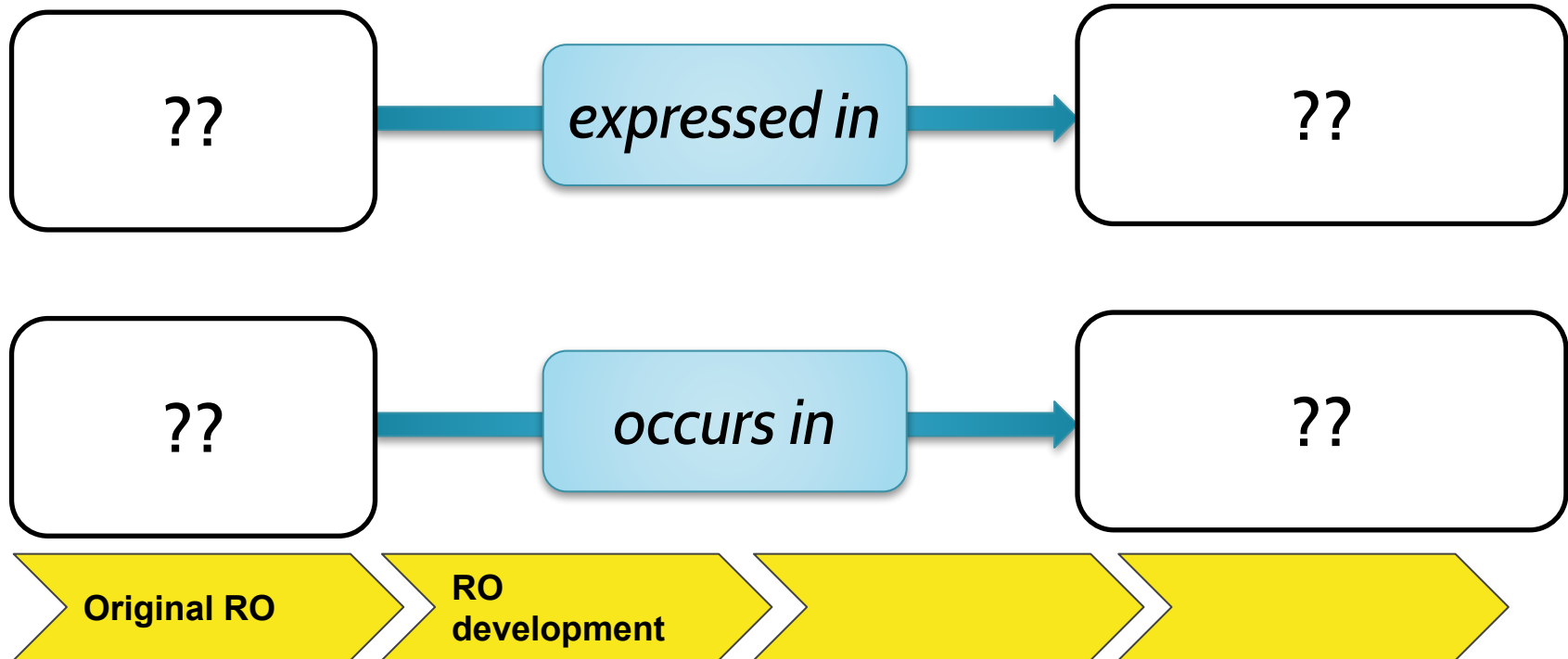


<https://oborel.github.io/obo-relations/ro-core/>

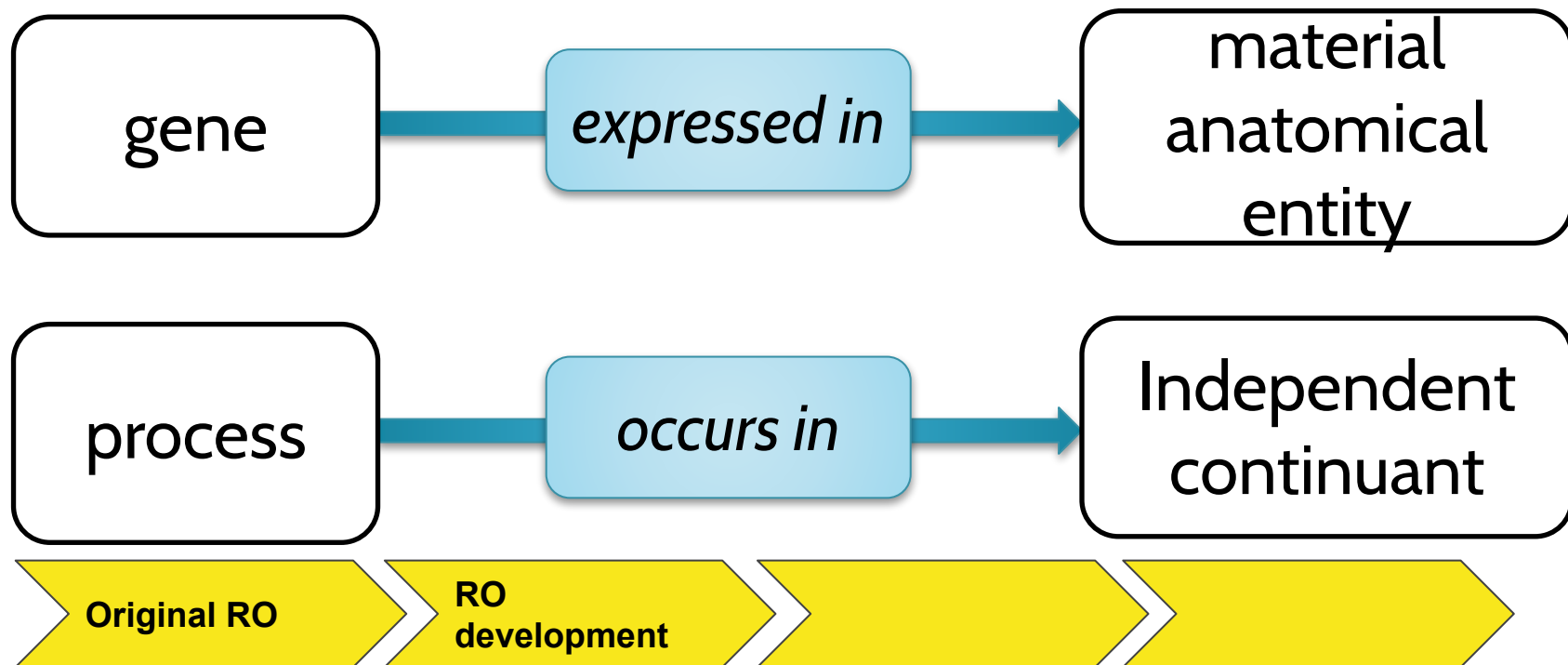
Original RO

RO
development

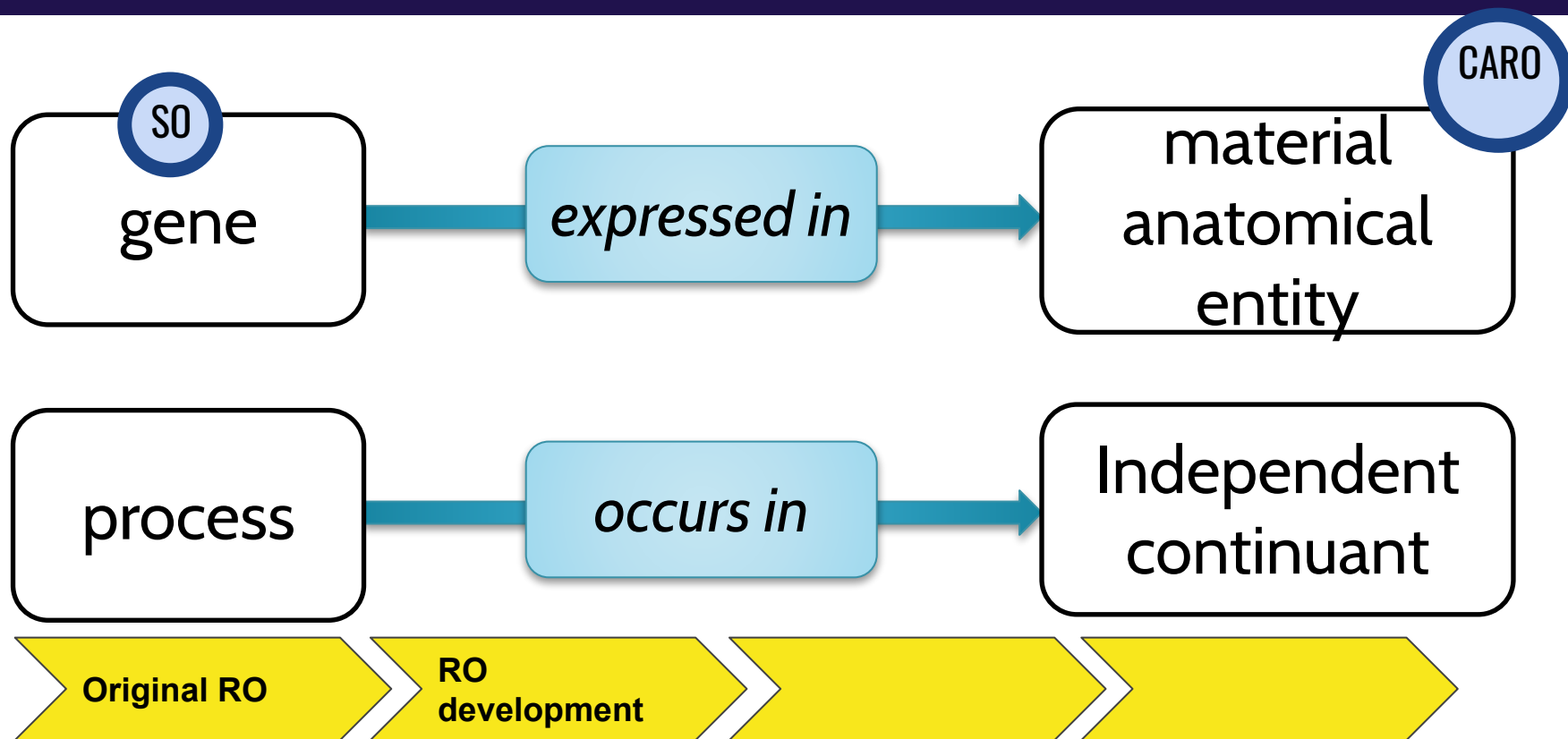
Axiomatizing relations using classes



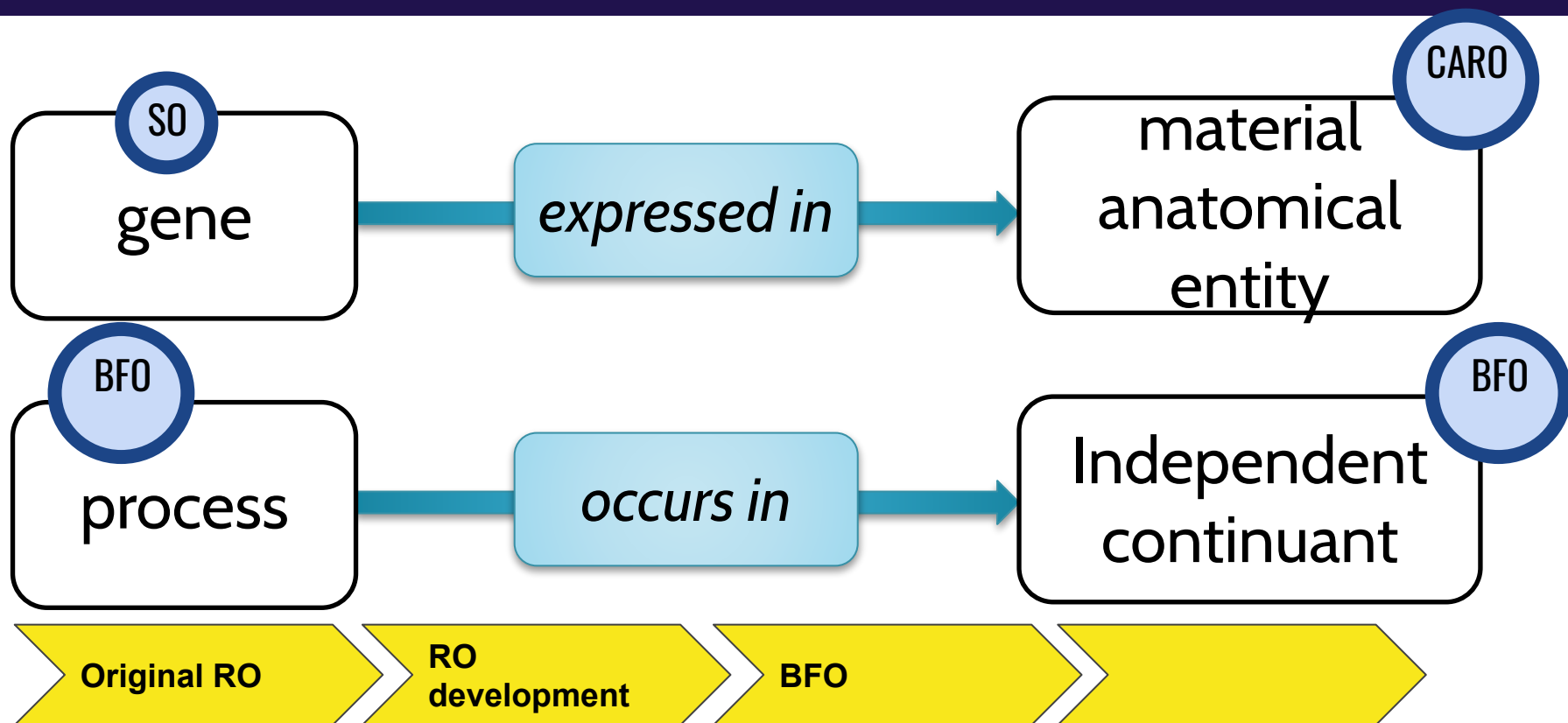
Axiomatizing relations using classes



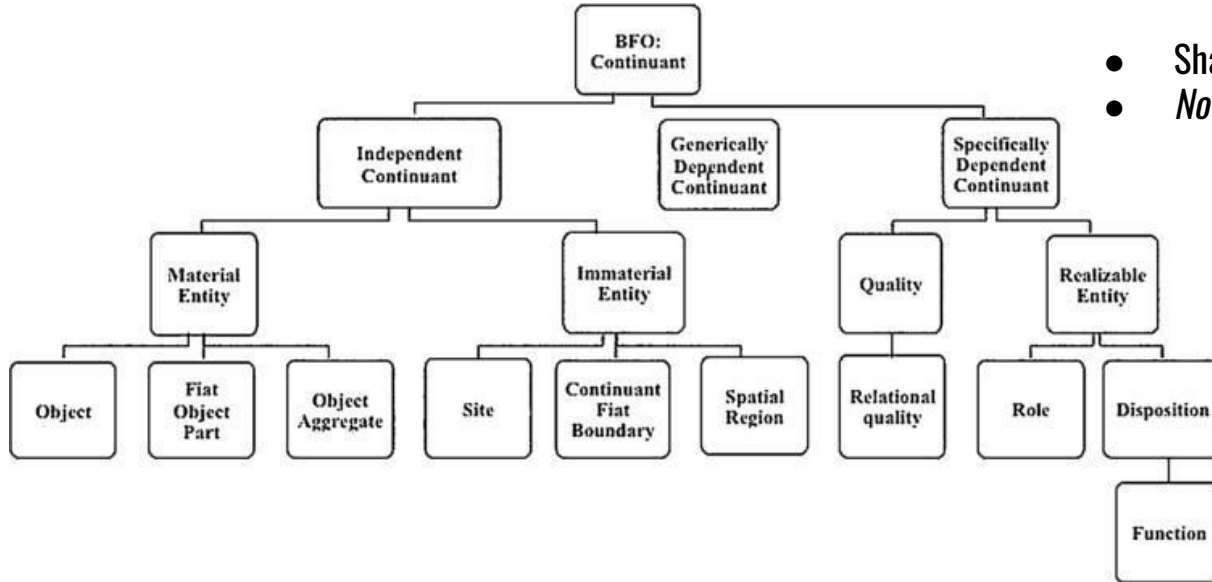
Root nodes of domain OBOs



Abstract relations and classes



Basic Formal Ontology



- Shared philosophical abstractions
- *Not* intended as a biological top layer

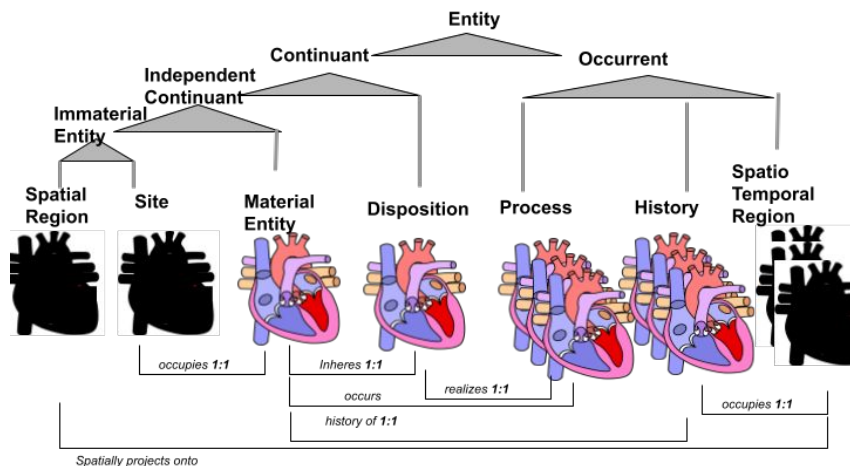


Original RO

RO
development

BFO

Multiple abstract aspects of concept



The lack of a consensus *biological* top level has led to duplicative efforts across ontologies in different branches



Just tell me
where to place
my anatomy
terms

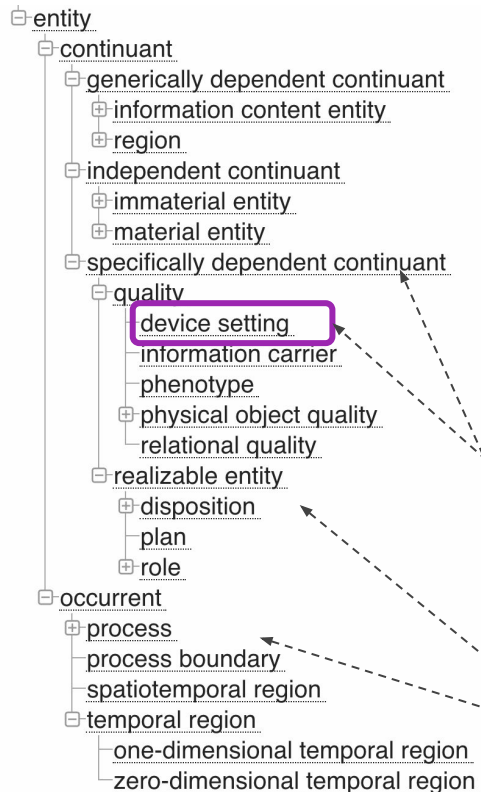
<https://douroucouli.wordpress.com/2022/08/10/shadow-concepts-considered-harmful/>

Original RO

RO
development

BFO

Problem: leaky abstractions



Ontology for Biomedical Investigations

The Ontology for Biomedical Investigations (OBI) is build in a collaborative, international effort and will serv instrumentation used, the data generated and the types of analysis performed on the data. This ontology ar common to all biomedical investigations, including functional genomics investigations and those that are mc

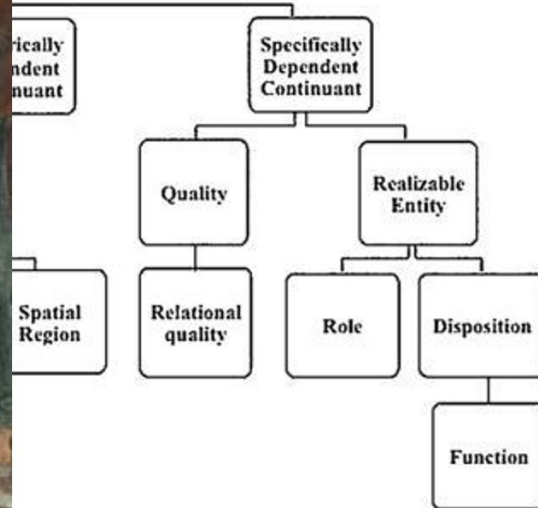
*OBI
term*

*Upper ontology
terms*

I just want to
explore assay
terms



Problem: level of abstraction is too high



I just want to
explore assay
terms



Desiderata for a biological upper core

- should provide a **parent to every OBO ontology class**
- should be anchored in BFO, but **hide its complexity from end-users**
- should include **logical axioms** that make inconsistencies within and between ontologies apparent through reasoning



COB: Common Ontology for Biology

Core Ontology for Biology and Biomedicine

COB brings together key terms from a wide range of OBO projects to improve interoperability.

[OntoBee](#)[AberOWL](#)[OLS](#)[Bioregistry](#)

The Core Ontology for Biology and Biomedicine (COB) brings together key terms from a wide range of OBO projects into a single, small ontology. The goal is to improve interoperability and reuse across the OBO community through better coordination of key terms.

Products

cob.owl	COB	Core Ontology for Biology and Biomedicine, main ontology
cob/cob-base.owl	COB base module	base module for COB
cob/cob-to-external.owl	COB to external	
cob/products/demo-cob.owl	COB demo ontology (experimental)	demo of COB including subsets of other ontologies (Experimental, for demo purposes only)

Original RO

RO
development

BFO

COB

ID Space[cob](#)**PURL**<http://purl.obolibrary.org/obo/cob.owl>**License**[CC0 1.0](#)**Homepage**<https://github.com/OBOFoundry/COB>**Contact**[Bjoern Peters](#)**Tracker**<https://github.com/OBOFoundry/COB/issues>**Domain**[upper](#)**Stars**

stars 24

Contributors

contributors 8

Last Commit

last commit august

[View](#)[Edit](#)[PURL](#)

Generated by: [_layouts/ontology_detail.html](#)

See [metadata guide](#)

Documentation

🏠 Core Ontology for Biology and Biomedicine

Search docs

Getting started

TOUR

Exploring COB

Placing OBO ontologies under COB

Using COB

Integration Tests

HOW-TO GUIDES

Standard ODK workflows

Overview

Editors Workflow

Release Workflow

Manage your ODK Repository

Setting up Docker for ODK

Imports management

Managing the documentation

Continuous Integration

Your ODK Repository Overview

Contributing

OTHER

Cite

[Docs](#) » Getting started

COB Ontology Documentation

This project is an attempt to bring together key terms from a wide range of [Open Biological and Biomedical Ontology \(OBO\)](#) projects into a single, small ontology. The goal is to improve interoperability and reuse across the OBO community through better coordination of key terms. Our plan is to keep this ontology small, but ensure that one or more COB terms can be used as the root of any given OBO library ontology.

Editors Guide

You can find descriptions of the standard ontology engineering workflows [here](#).

Issue Tracker

If you are responsible for an OBO ontology you can see any tickets that pertain to your ontology by looking for the label with your ontology ID.

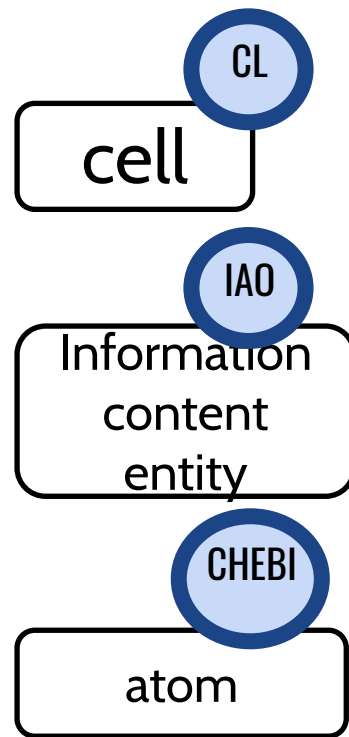
For example:

- **GO:** <https://github.com/OBOFoundry/COB/labels/GO>
- **OBI:** <https://github.com/OBOFoundry/COB/labels/OBI>

Next ➞

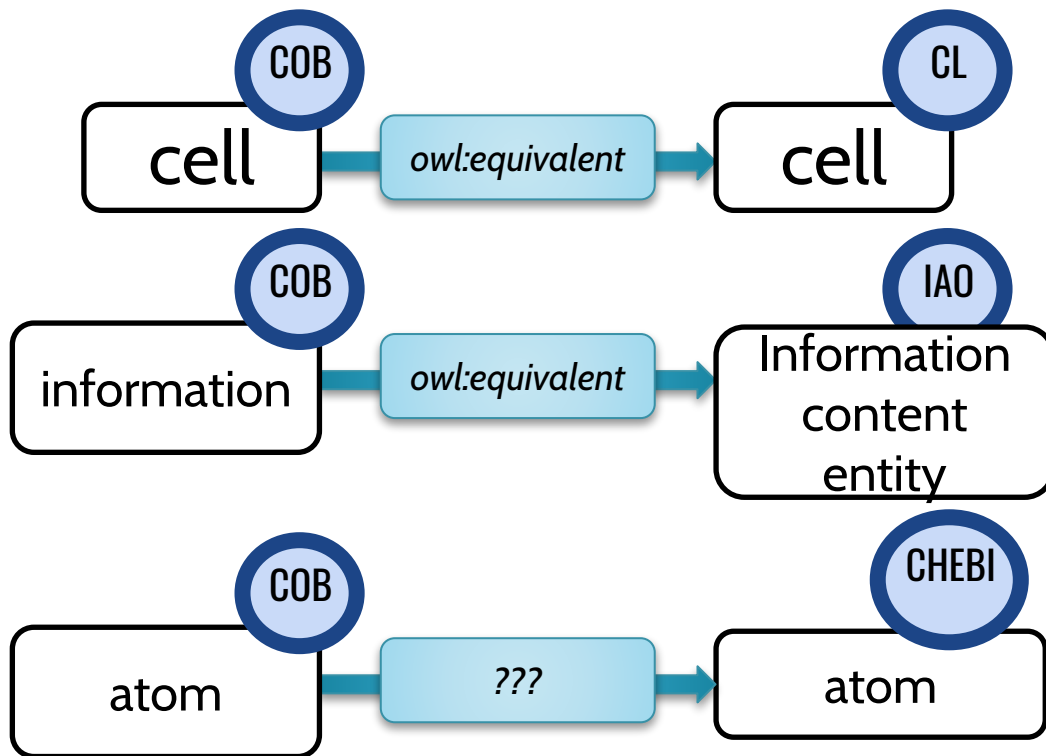
Bottom-up approach to design

- Examine top level “native” terms in core OBO ontologies



COB and OBO top levels

- COB classes shadow existing OBO top levels
- Semantics may vary
 - If semantics are identical, we retain equivalence mappings
 - Label changes OK



COB and OBO top levels

- COB classes shadow existing OBO top levels
- Semantics may vary
 - If semantics are identical, we retain equivalence mappings

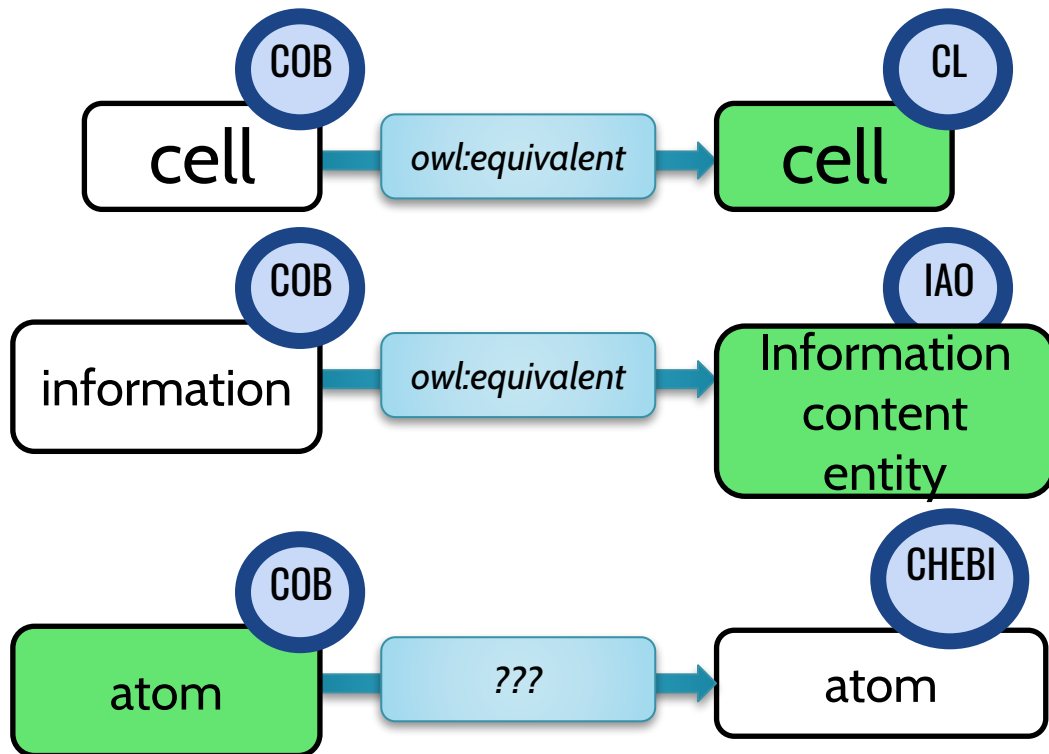
Example entries

The following entries use the owl:equivalentClass predicate to indicate that the COB IDs are equivalent to the the corresponding OBO concept.

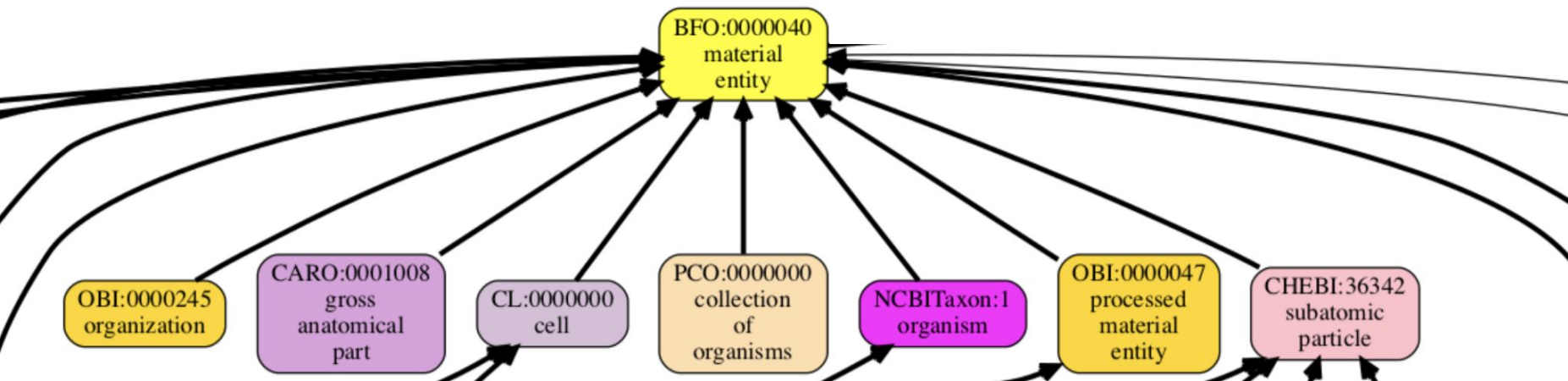
subject_id	subject_label	predicate_id	object_id	object_label
COB:0000003	mass	owl:equivalentClass	PATO:0000125	mass
COB:0000004	charge	owl:equivalentClass	PATO:0002193	electric
COB:0000006	material entity	owl:equivalentClass	BFO:0000040	material entit
COB:0000008	proton	owl:equivalentClass	CHEBI:24636	proton
COB:0000017	cell	owl:equivalentClass	CL:0000000	cell
COB:0000026	processed material entity	owl:equivalentClass	OBI:0000047	processed ma
COB:0000031	immaterial entity	owl:equivalentClass	BFO:0000041	immaterial er
COB:0000033	realizable	owl:equivalentClass	BFO:0000017	realizable ent

Existing OBO IDs are used in release version

- ID swapped out *when there is an equivalent*



Existing OBO IDs are used in release version



<http://purl.obolibrary.org/obo/cob.owl>

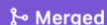
COB development process

- Workshops
 - 2018 RO meeting (Denver)
 - 2019 ICBO (IRL)
 - 2020 ICBO (Virtual)
 - 2021 ICBO (Virtual)
- Slack channel
- Modern GitHub-based workflows
 - See ODK talk Feb 8
- Goals:
 - No behind-doors decision making
 - All decisions transparent on issue tracker
 - Anyone can make a PR

Practicality, use cases >>
philosophy, perfectionism

Open: community can make PRs

Add superclass mapping from COB characteristic to ChEBI role #173

[Edit](#)[Open with ▾](#)**Merged**

cmungall merged 7 commits into `0B0Foundry:master` from `cthoyt:align-chemi-role` 5 hours ago

[Conversation](#) 6[Commits](#) 7[Checks](#) 1[Files changed](#) 2[+3 -1](#)

cthoyt commented 2 days ago • edited ▾

[Contributor](#)

Closes [#169](#)

This PR adds an explicit equivalence mapping between COB and ChEBI:

- COB:0000502 (characteristic) `sssom:superClassOf` CHEBI:50906 (role)
- COB:0000088 (drug product) `rdfs:seeAlso` CHEBI:23888 (drug)

This one seems pretty straightforward, so maybe there's a reason these were difficult to align before, so feedback or changes (e.g., different relationship) are welcome.

Note: I tried running the `make` command but got an error (see [#174](#)). I'm curious if there are tools built in to the `make` command that can automatically check that adding an equivalence like this doesn't cause any problems

Pipeline

No Workspace yet - [Create One](#)

Reviewers



cmungall



Assignees



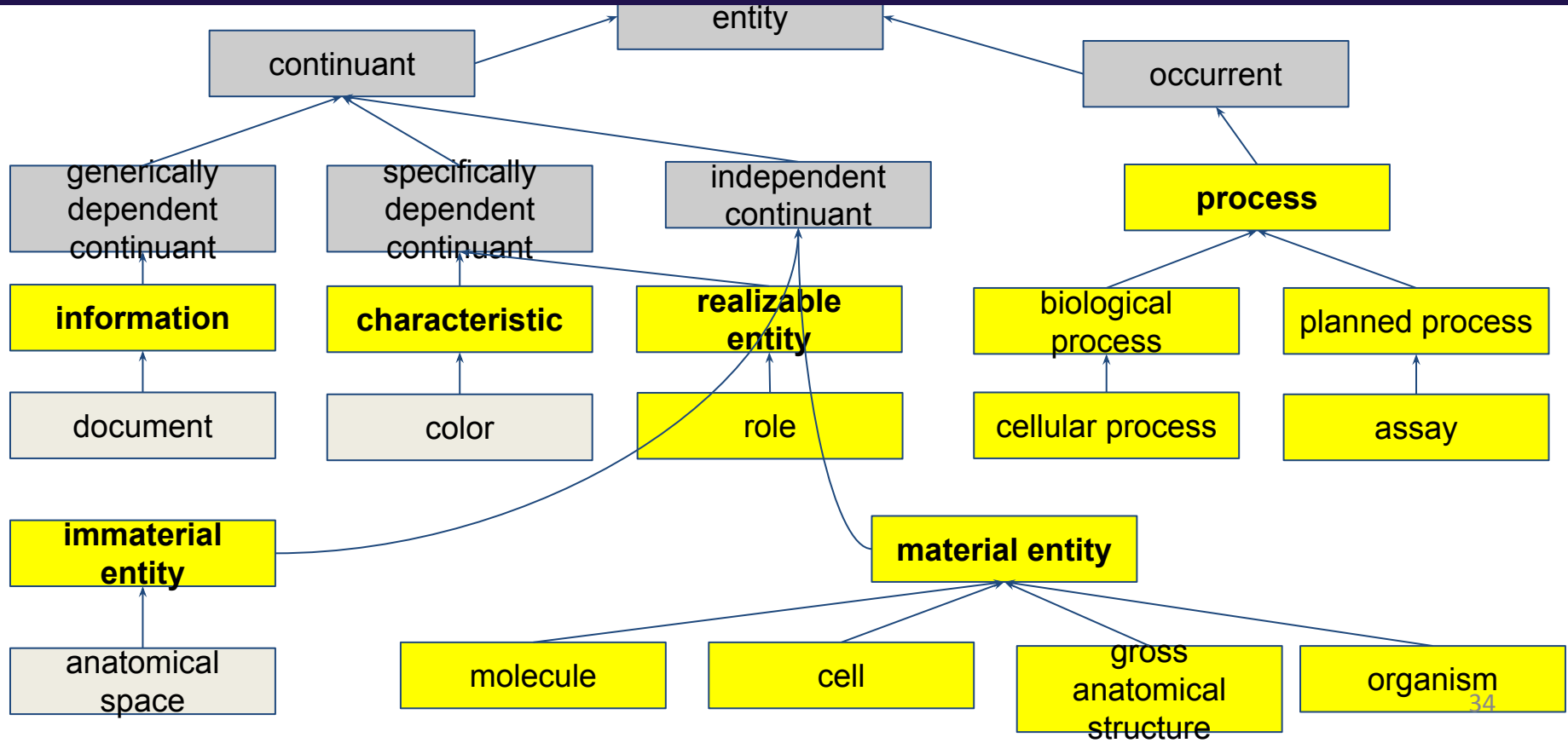
No one—assign yourself

Labels



None yet

COB and BFO



Flattened view of BFO for COB

process

material entity

site

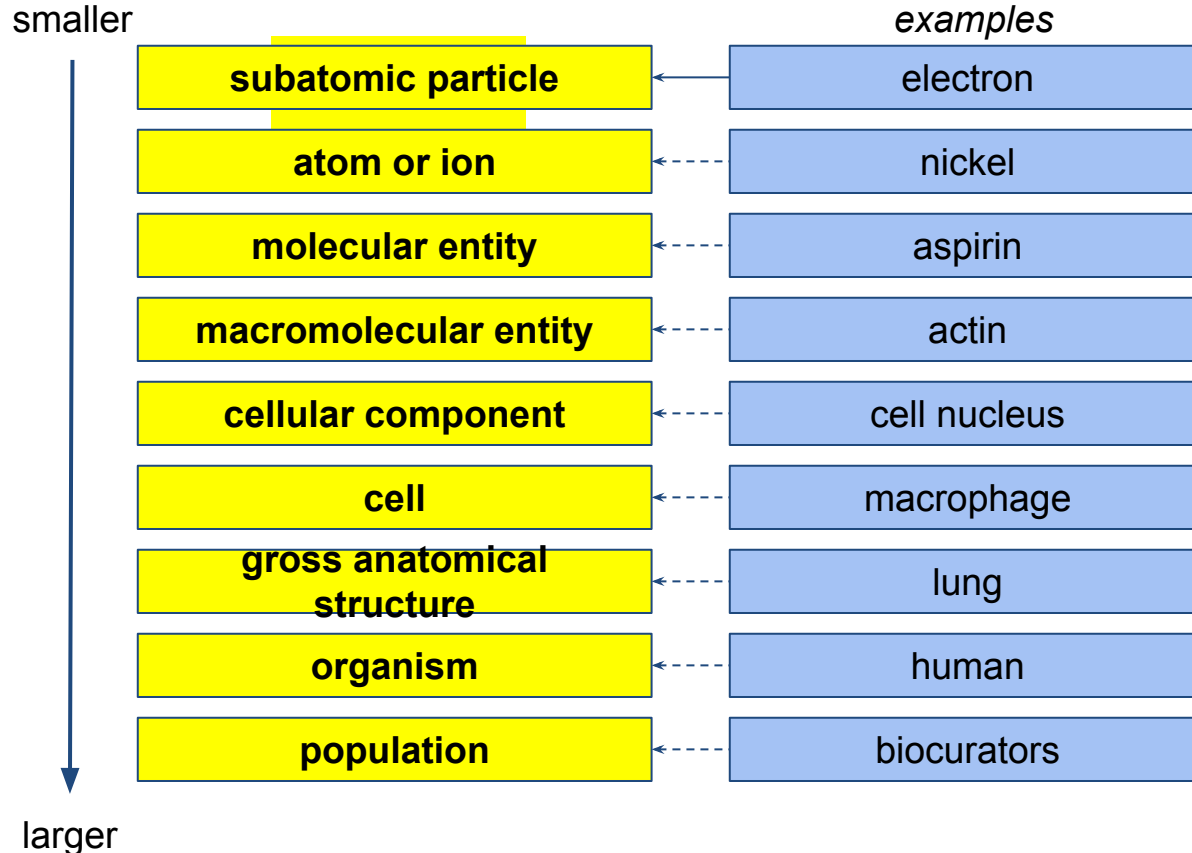
characteris
tic

realizable entity

information

- Include only terms that have OBO subclasses
- Use bfo terms that are non-threatening for biologist ontology consumers
- Keep full logical axiomatization ‘in the background’

COB – Material Entities



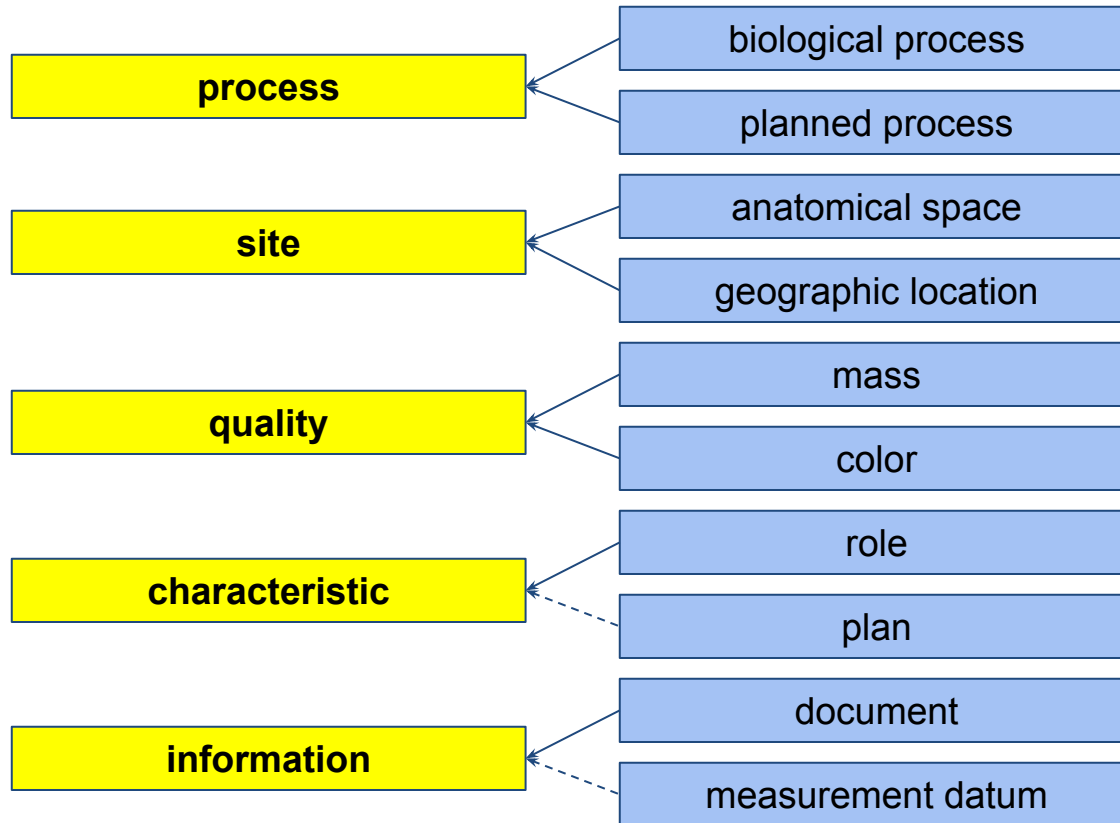
Generic: material entity

Physical world: subatomic particle, atom or ion, molecular entity, macromolecular entity

Biological: cellular component, cell, gross anatomical structure, organism, population

Human activities: processed material (not shown)

COB – Non-Material Entities



Generic: process, quality, realizable

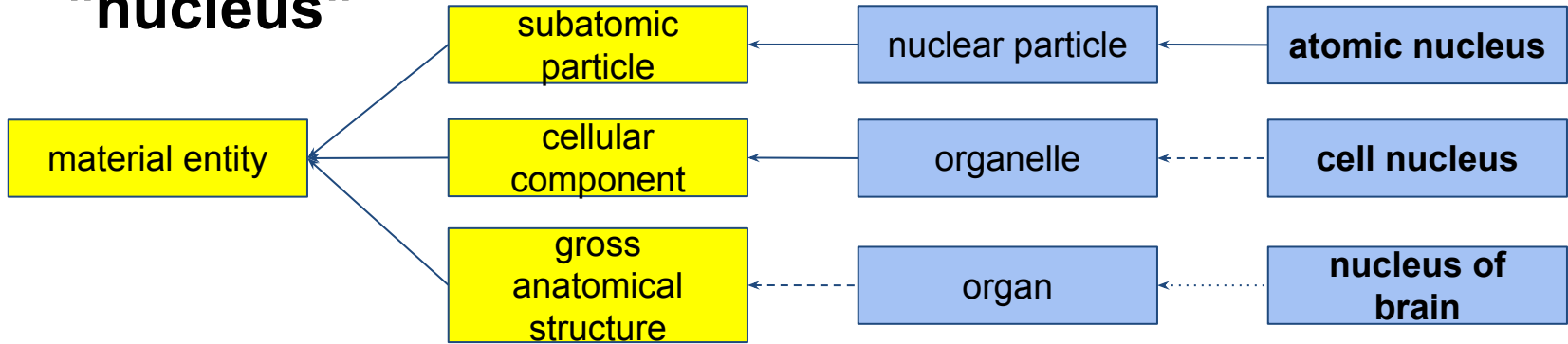
Physical world: site, geographic location, mass, color

Biological: biological process, anatomical space, role

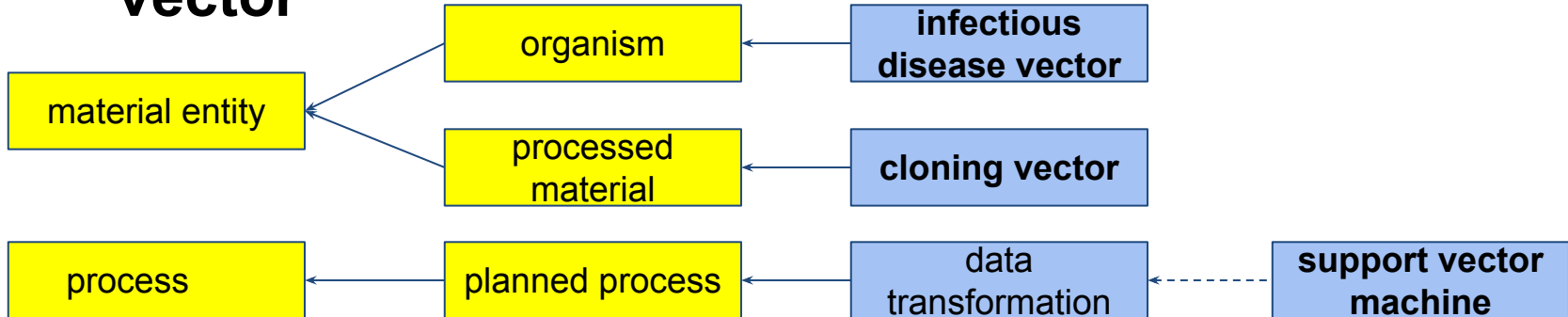
Human activities: planned process, information

COB – Examples

"nucleus"

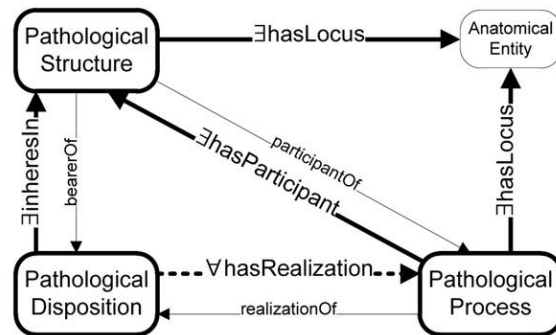


"vector"



In progress: phenotype and disease

- Upper ontology assumptions don't align with actual uses of phenotype and disease ontologies
- Consider Schulz conflation model?



[J Biomed Semantics](#), 2011; 2(Suppl 2): S6.

Published online 2011 May 17. doi: [10.1186/2041-1480-2-S2-S6](#)

PMCID: PMC3102895

PMID: [21624161](#)

Scalable representations of diseases in biomedical ontologies

[Stefan Schulz](#),^{1,5} [Kent Spackman](#),² [Andrew James](#),³ [Cristian Cocos](#),⁴ and [Martin Boeker](#)⁵

[Author information](#) [Article notes](#) [Copyright and License information](#) [Disclaimer](#)

This article has been [cited by](#) other articles in PMC.

Abstract

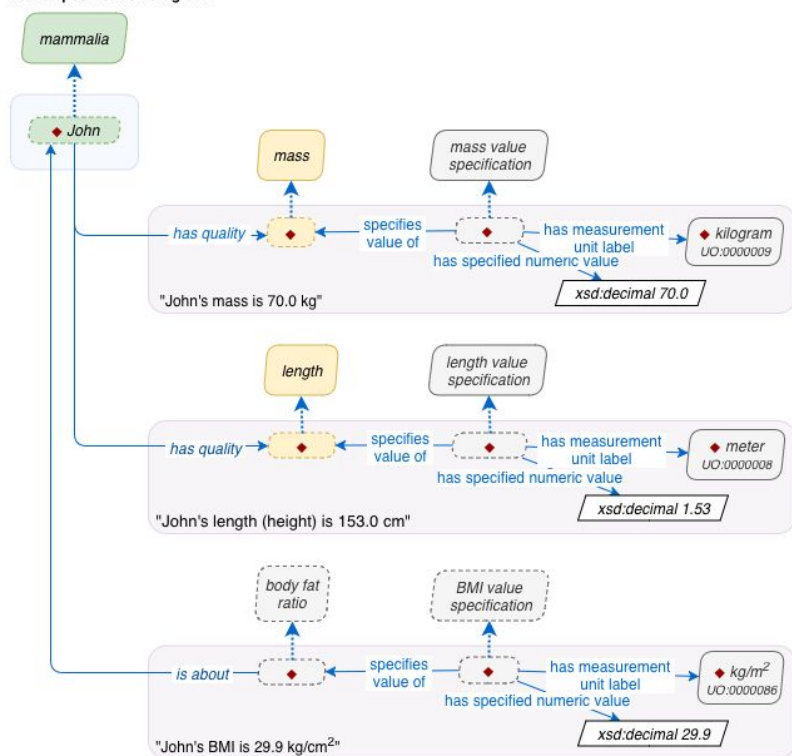
[Go to:](#) [📄](#)

Background

The realm of pathological entities can be subdivided into pathological dispositions, pathological processes, and pathological structures. The latter are the bearer of dispositions, which can then be realized by their manifestations — pathologic processes. Despite its ontological soundness, implementing this model via purpose-oriented domain ontologies will likely require considerable effort, both in ontology construction and maintenance, which constitutes a considerable problem for SNOMED CT, presently the largest biomedical ontology.

In progress: units and measurements

Value specification diagram



Proposed
simplification of
existing
measurement model

Alignment with Biolink

BioPortal Ontologies Search Annotator Recommender Mappings

Biolink Model
Last uploaded: April 12, 2021

Summary Classes Properties Notes Mappings Widgets

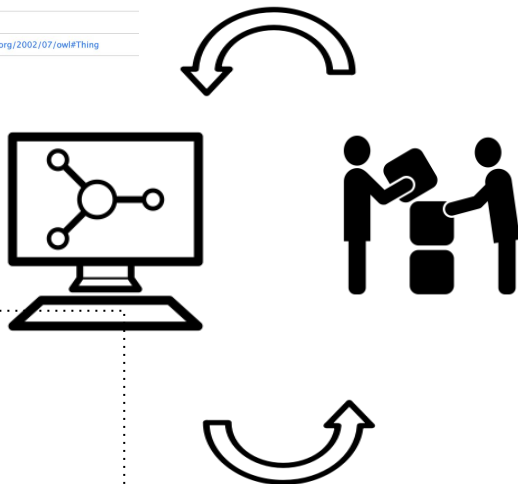
Jump to:

- annotation
 - boolean
 - chemical formula value
 - class_definition
 - date
 - datetime
 - decimal
 - double
 - entity
 - association
 - named thing
 - activity
 - administrative entity
 - biological entity
 - biological process or activity
 - disease or phenotypic feature
 - disease
 - disease or phenotypic feature exposure
 - disease or phenotypic feature outcome
 - phenotypic feature
 - epidemiological outcome
 - molecular entity
 - organismal entity
 - anatomical entity
 - cell
 - cellular component
 - gross anatomical structure
 - pathological anatomical structure
 - cell line
 - individual organism
 - life stage
 - population of individual organisms

Preferred Name	annotation
Definitions	Biolink Model root class for entity annotations.
ID	https://w3id.org/biolink/vocab/Annotation
definition	Biolink Model root class for entity annotations.
label	annotation
prefLabel	annotation
subClassOf	http://www.w3.org/2002/07/owl#Thing

Developed as part of NCATS Translator project

- **Weekly Data Modeling calls** (20-40 people)
- **Working groups** for different areas (e.g. chemicals)
- **Technically diverse group** (domain scientists, bioinformaticians, ontologists)
- **Use of GitHub** (PRs, votes)



Biolink-Model: A schema for biological KGs

- Expressed in LinkML
- "Ontology-like"
- OBO classes are instances
- Edges are first-class

<https://biolink.github.io/biolink-model>

Conclusions: COB...

- ...provides a common biological upper level for OBO
- ...is composed largely from roots of existing OBO ontologies
- ...is compatible with yet insulates from BFO
- ...is fully open and welcomes all contributions

Resources



COB on OBO:

<https://obofoundry.org/ontology/cob/>

Website:

<https://obofoundry.org/COB/>

GitHub:

<https://github.com/OBOFoundry/COB>

— — —

Acknowledgements

— — —

Bjoern Peters

James Overton

Charles Hoyt

Nico Matentzoglou

Anne Thessen

Damion Dooley

Chris Stoeckert

Duncan, Bill

Natale, Darren

Nomi Harris

Rebecca Jackson

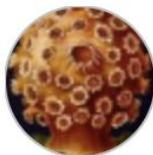
Randi Vita

Yongqun He

Melissa Haendel

Barry Smith

NIH R24HG010032



<https://contrib.rocks/preview?repo=OBOFoundry%2FCOB>