# COB: A Core Ontology for Biology

Ontology Summit 2023
https://ontologforum.org/index.php/OntologySummit2023

Chris Mungall

https://genomic.social/@cmungall
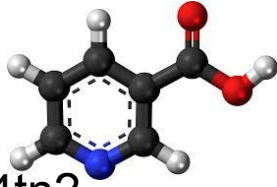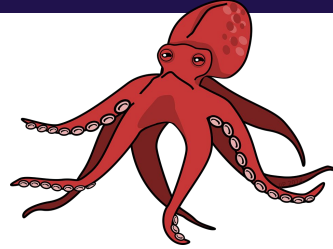
@chrismungall

Lawrence Berkeley National Laboratory

# We have many identifiable things and categories
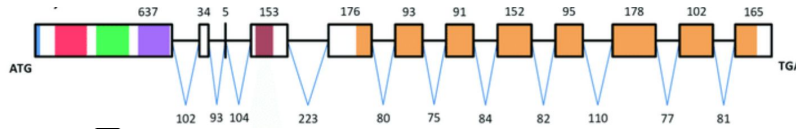


Drugs 10k
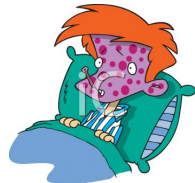Chemicals 4tn?

Species
~9 million

Cells
10,000s+
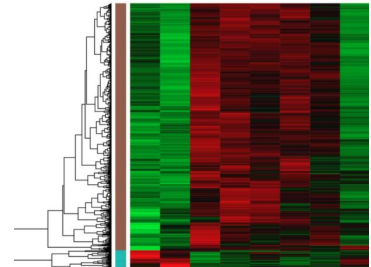types
per species)

Genes 20k per species
Metagenome dbs: **>65bn** genes
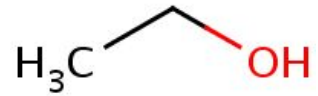
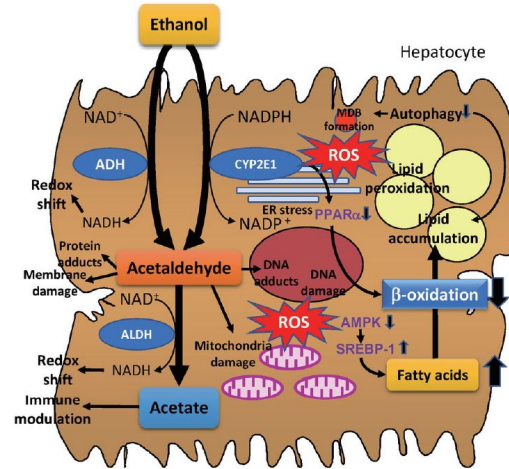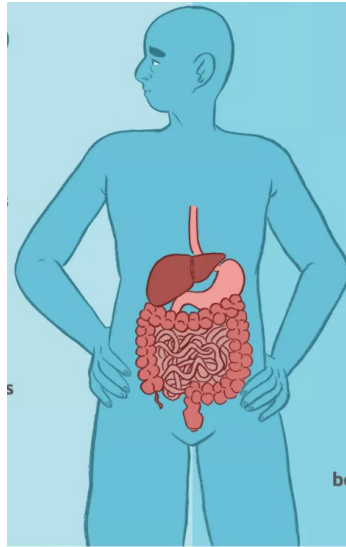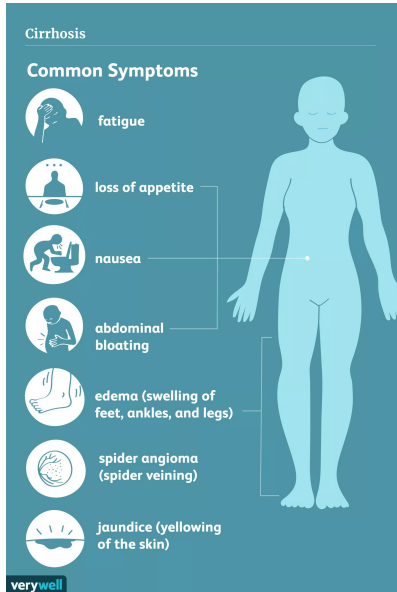Genetic
variants
3m in human
alone

Diseases and
Phenotypes
10-50k/species

Experiments
Raw data
?? exabytes

# The things are interconnected across scales

# Different perspectives on the same things
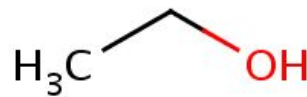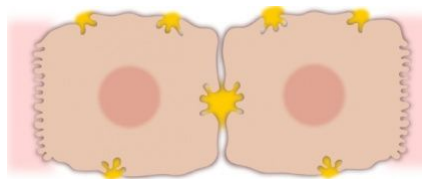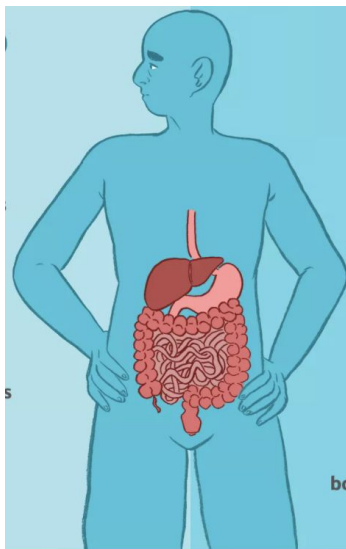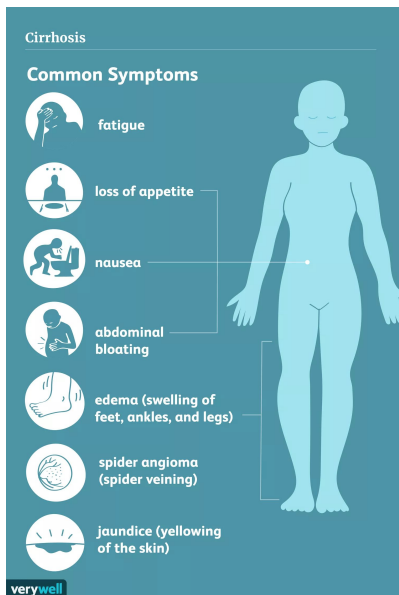


static | disease state



static | anatomy



static | cell structures



static | chem structures

# Different perspectives on the same things



dynamic | pathophysiology

dynamic | physiology

dynamic | pathways

dynamic | reactions
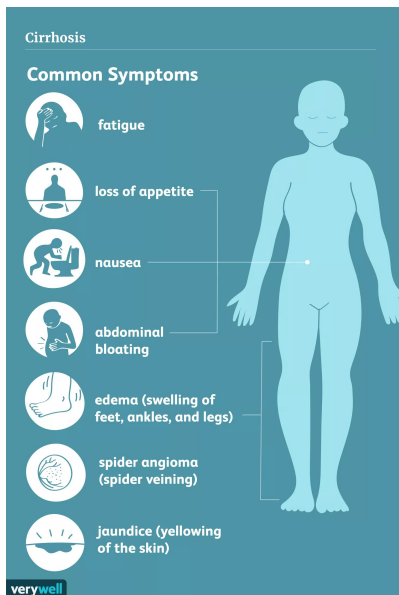
↑bilirubin → jaundice
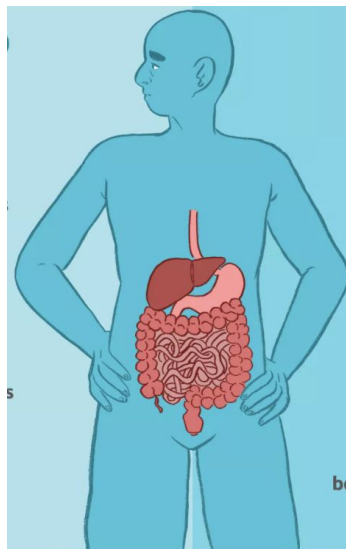
Conjugation of bilirubin occurs-in liver

$$CH_3CH_2OH \rightleftharpoons CH_3CHO$$
ethanol → acetaldehyde
$$NAD^+ \quad NADH+H^+$$
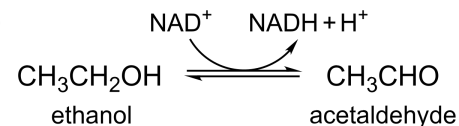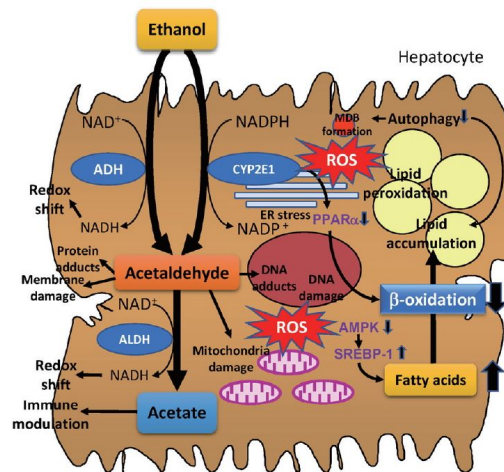
# OBO is for organizing the things

1. Well-integrated **Modular** ontologies
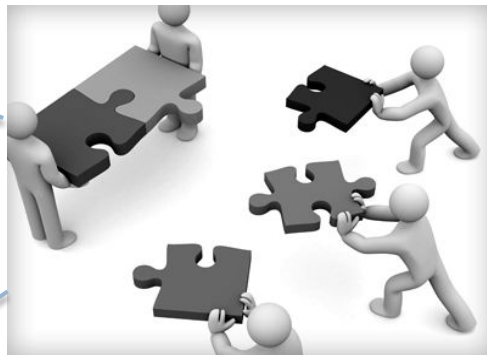
2. **Technical and sociotechnological framework** for cooperation

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

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



**F**indable **A**ccessible **I**nteroperable **R**eusable

@obofoundry

http://obofoundry.org

# OBO registry



https://obofoundry.org

# OBO ontologies *interoperate*
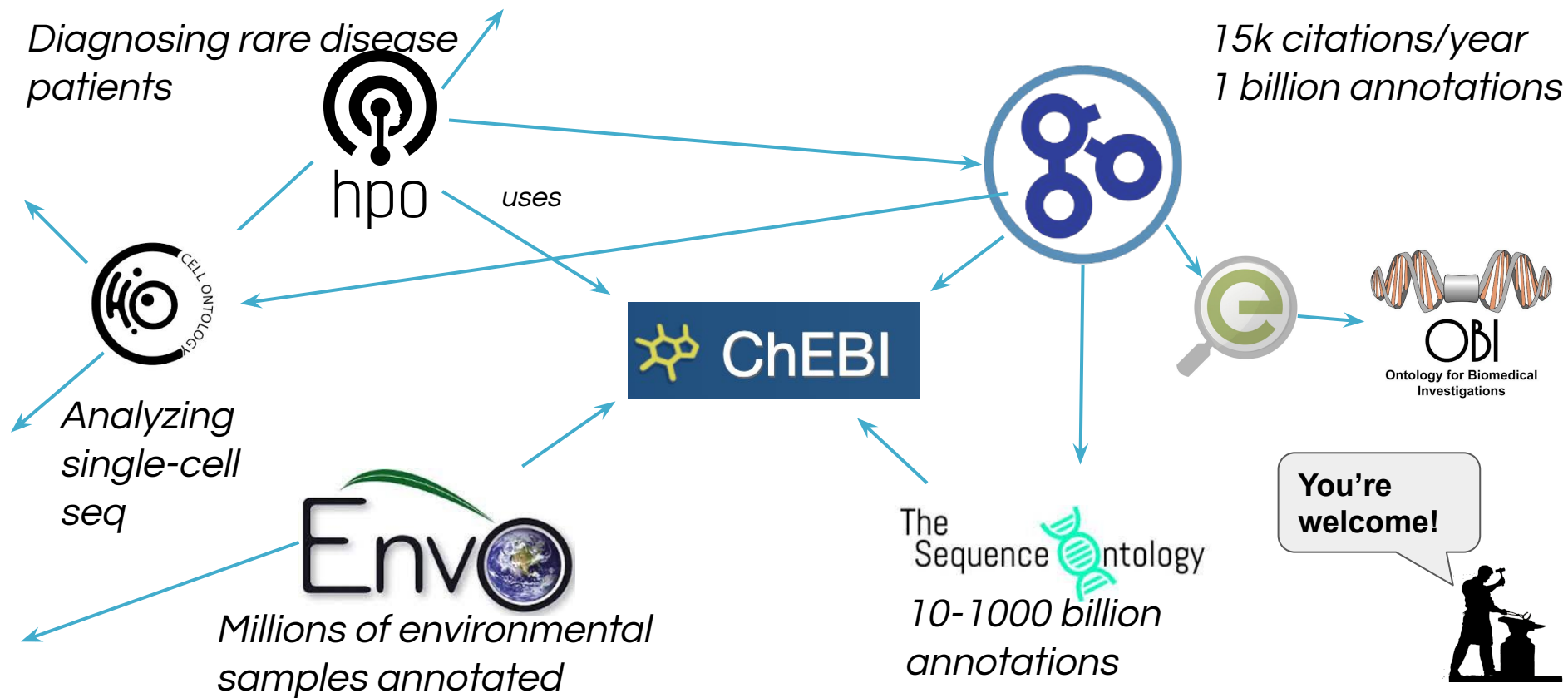
# Layers of OBO interoperability



**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, rdflib, fastobo, owltools, dosdp-tools etc are able to read and write ontologies in commonly understood formats.

Credit: Nico Matentzoglu

# Layers of OBO interoperability



Today

Mar 8, 15: Modules + Patterns

**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.

Feb 1:OAK
Feb 22: Ubergraph

**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, rdflib, fastobo, owltools, dosdp-tools etc are able to read and write ontologies in commonly understood formats.

Feb 1: ROBOT/OAK
Feb 8, 15: Tools

Credit: Nico Matentzoglu

1
2
3
4

# 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

| Relation | Transitive | Symmetric | Reflexive |
|---|---|---|---|
| is_a | + | - | + |
| part_of | + | - | + |
| located_in | + | - | + |
| contained_in | - | - | - |
| adjacent_to | - | - | - |
| transformation_of | + | - | - |
| derives_ from | + | - | - |
| preceded_by | + | - | - |
| has_participant | - | - | - |
| has_agent | - | - | - |

## 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

## 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

**Original RO**

# OWL formulation of RO



https://oborel.github.io



Original RO → RO development

# RO Core and extensions
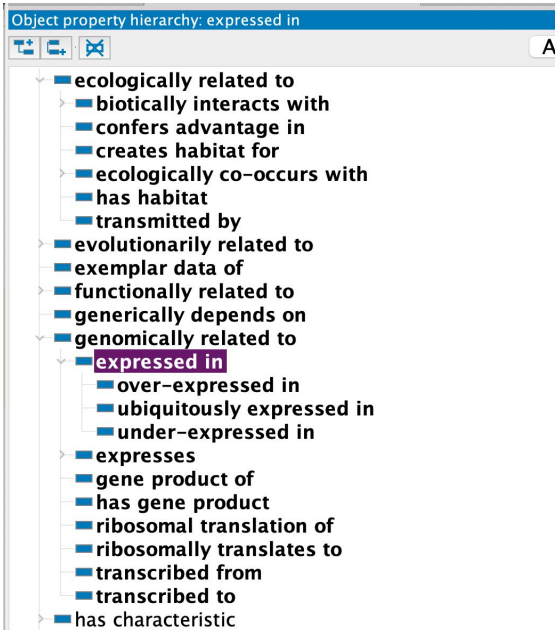


RO: **all** relations
- RO-Core: *cross-domain*
  - Part of
  - Located in
  - Occurs in
  - Characteristic of
- Ecological relations
  - Parasite of
  - ….
- Genomic relations
  - Expressed in
  - …
- …

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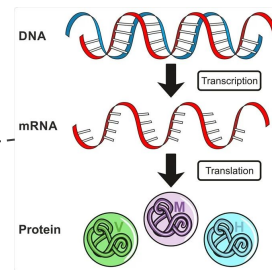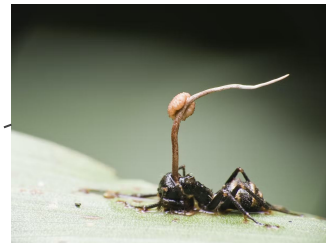




**Original RO** → **RO development** → →

# Axiomatizing relations using classes
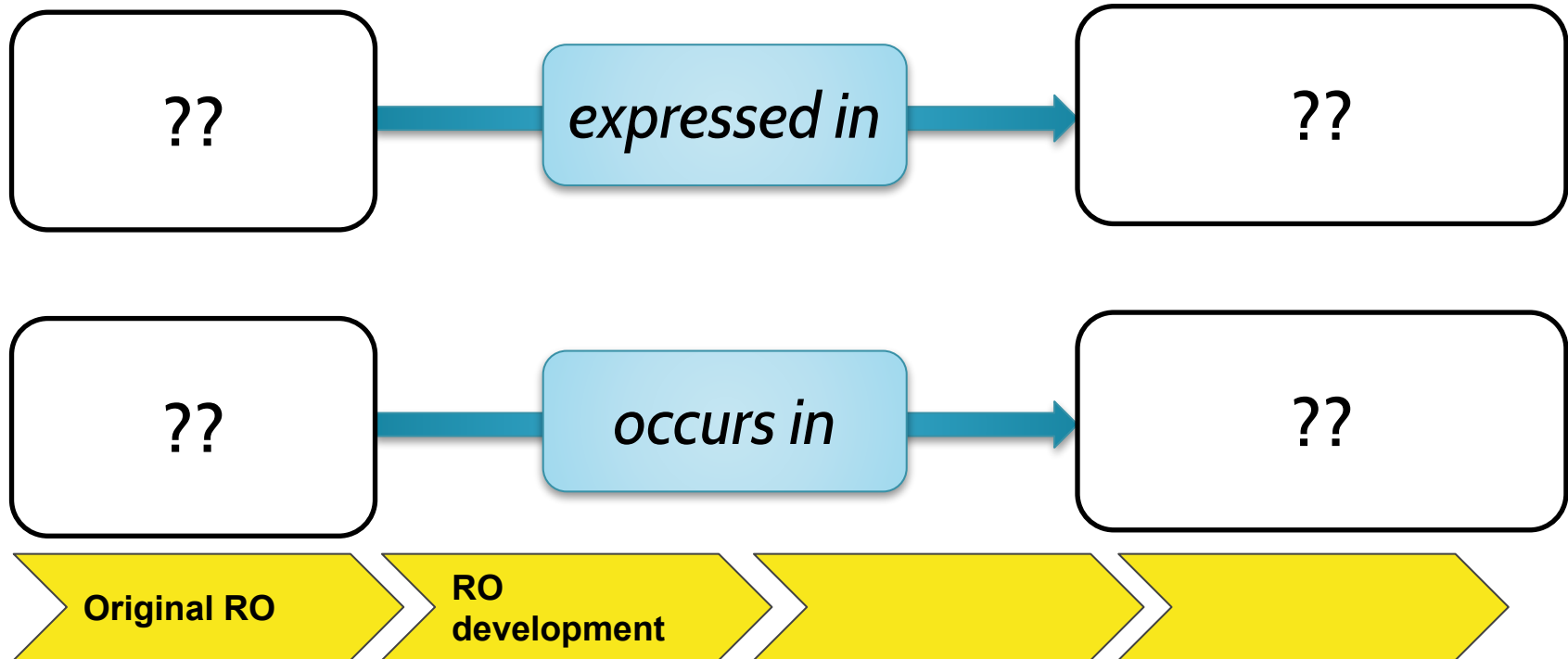
# Axiomatizing relations using classes
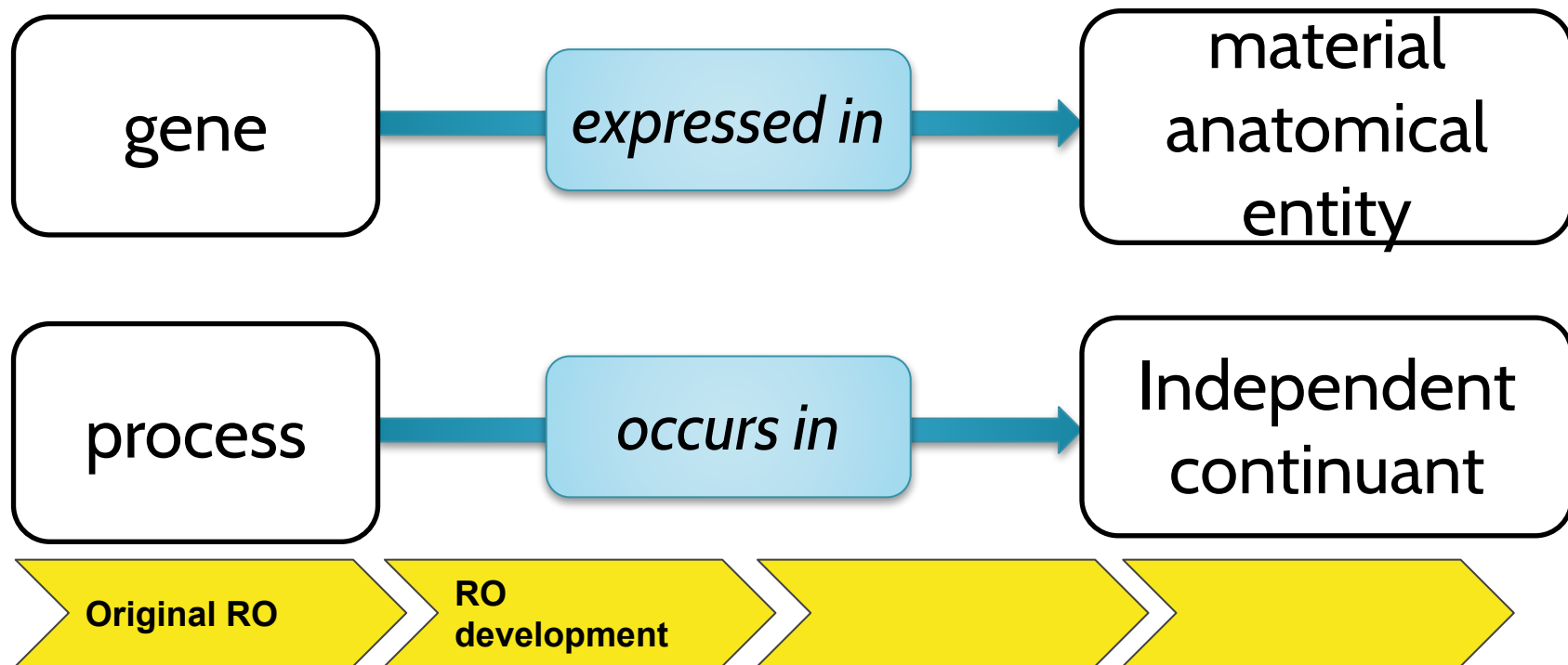
# Root nodes of domain OBOs
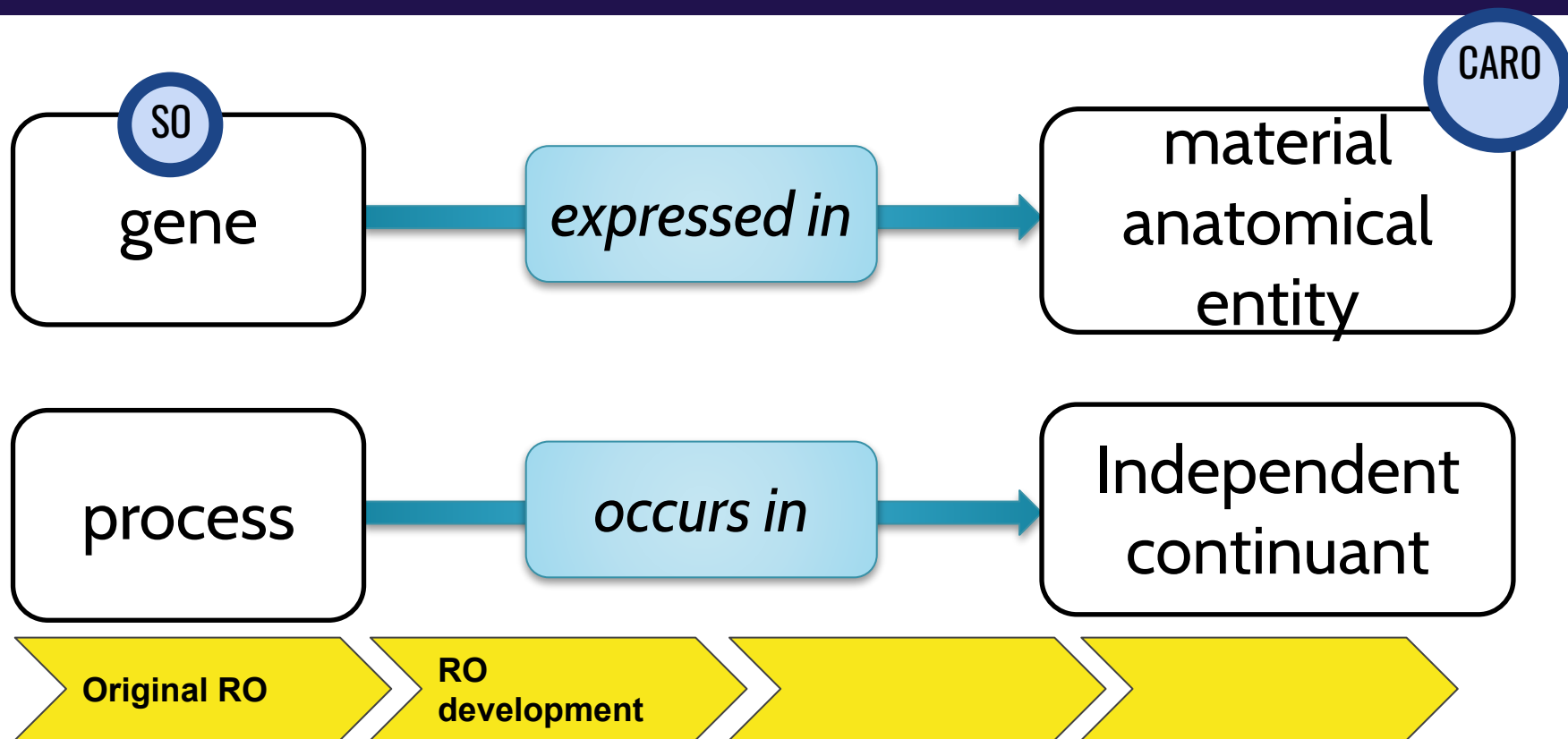
# Abstract relations and classes
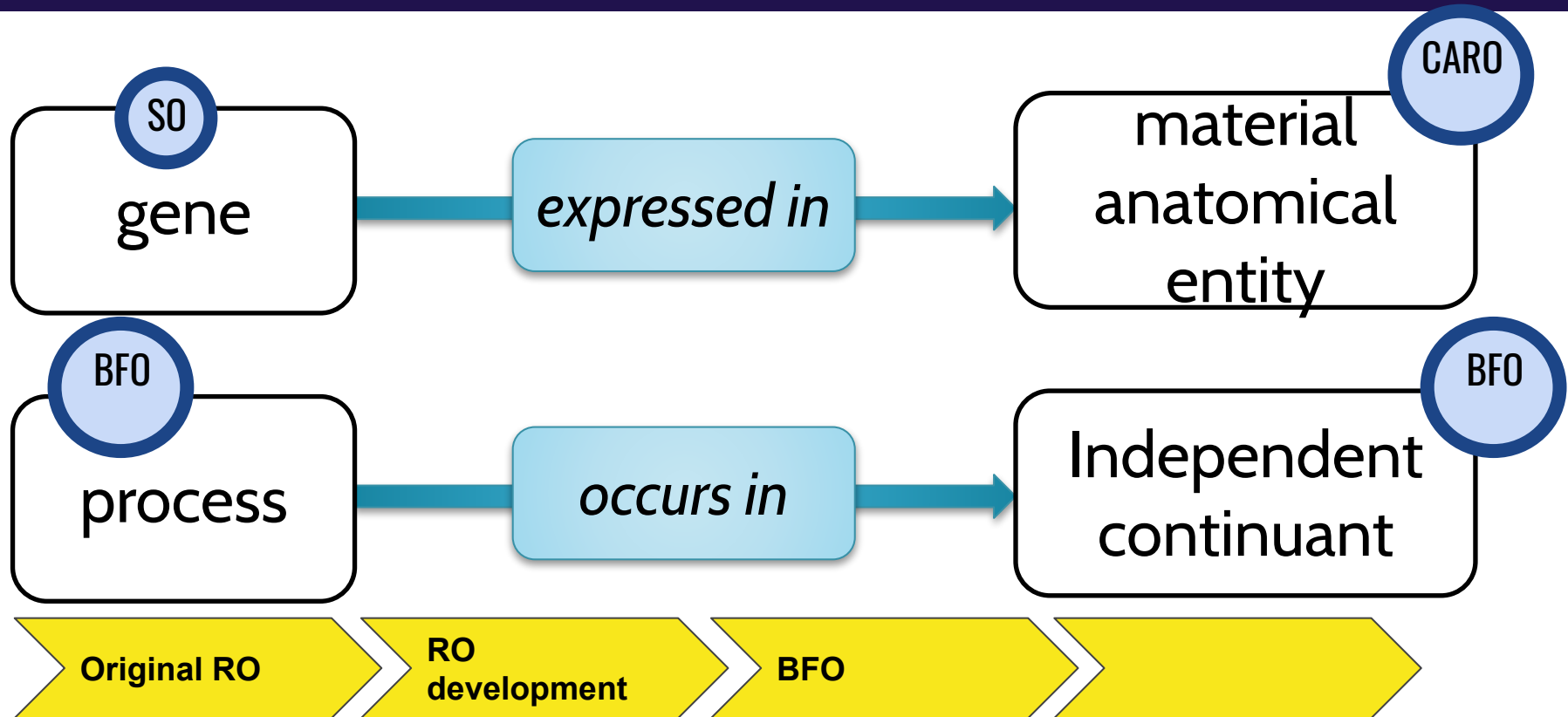
# Basic Formal Ontology
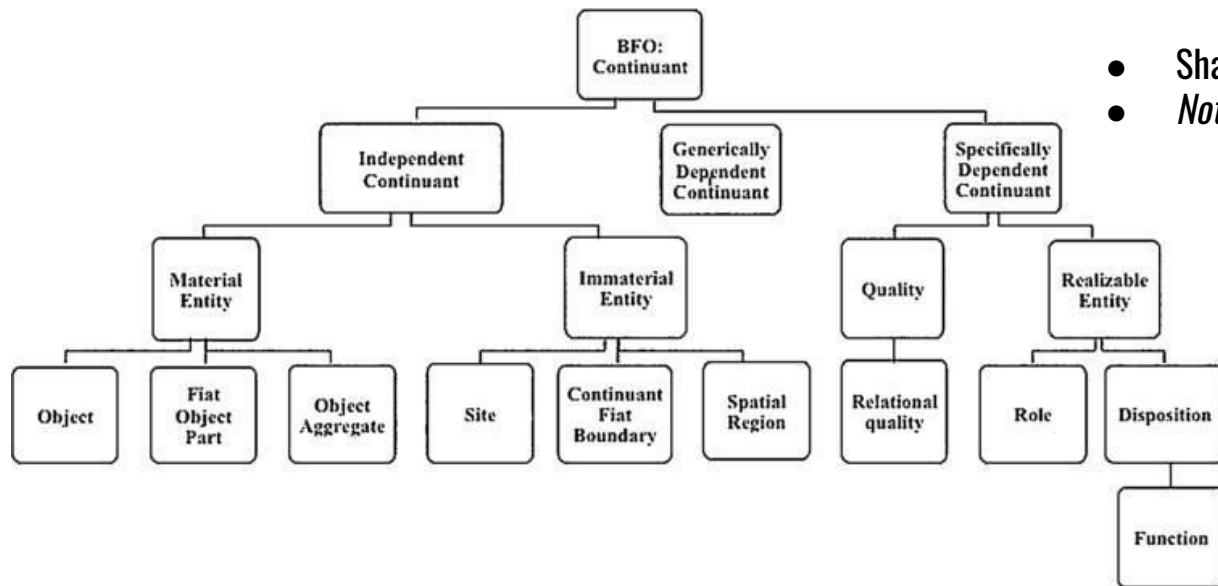


- Shared philosophical abstractions
- *Not* intended as a biological top layer



Original RO → RO development → BFO →

# Multiple abstract aspects of concept
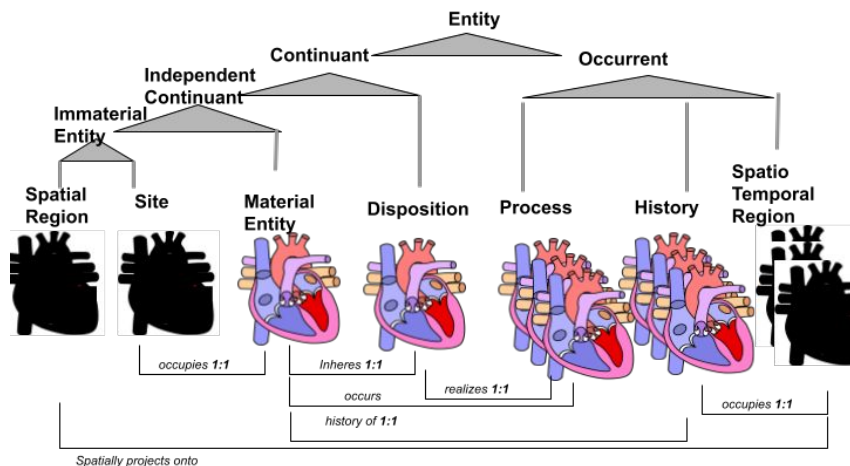


The lack of a consensus *biological* top level has led to duplicative efforts across ontologies in different branches

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
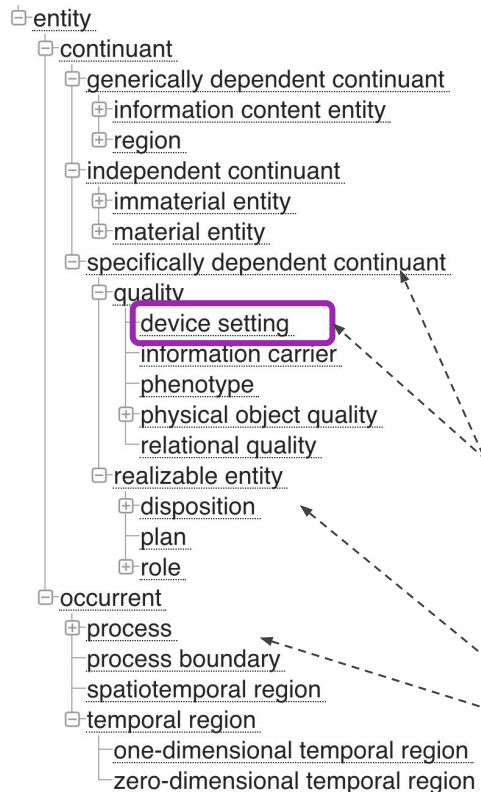
# Problem: level of abstraction is too high
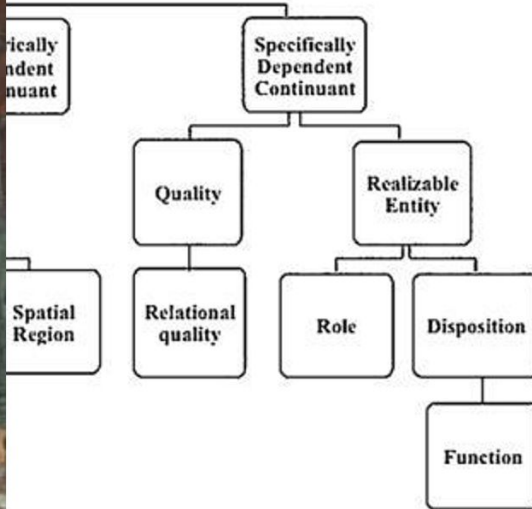
# Desiderata for a biological upper core

- 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

**Original RO** → **RO development** → **BFO** → **BFO**

# 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 interoperabilty 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) |

**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

**Original RO** → **RO development** → **BFO** → **COB**

# Documentation

# Bottom-up approach to design

- Examine top level "native" terms in core OBO ontologies
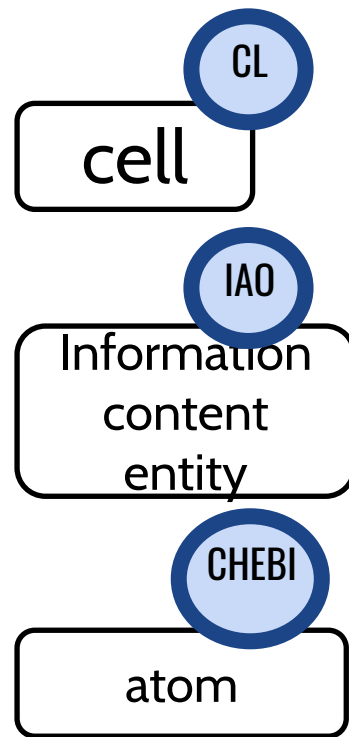
**CL** — cell

**IAO** — Information content entity

**CHEBI** — atom

# 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
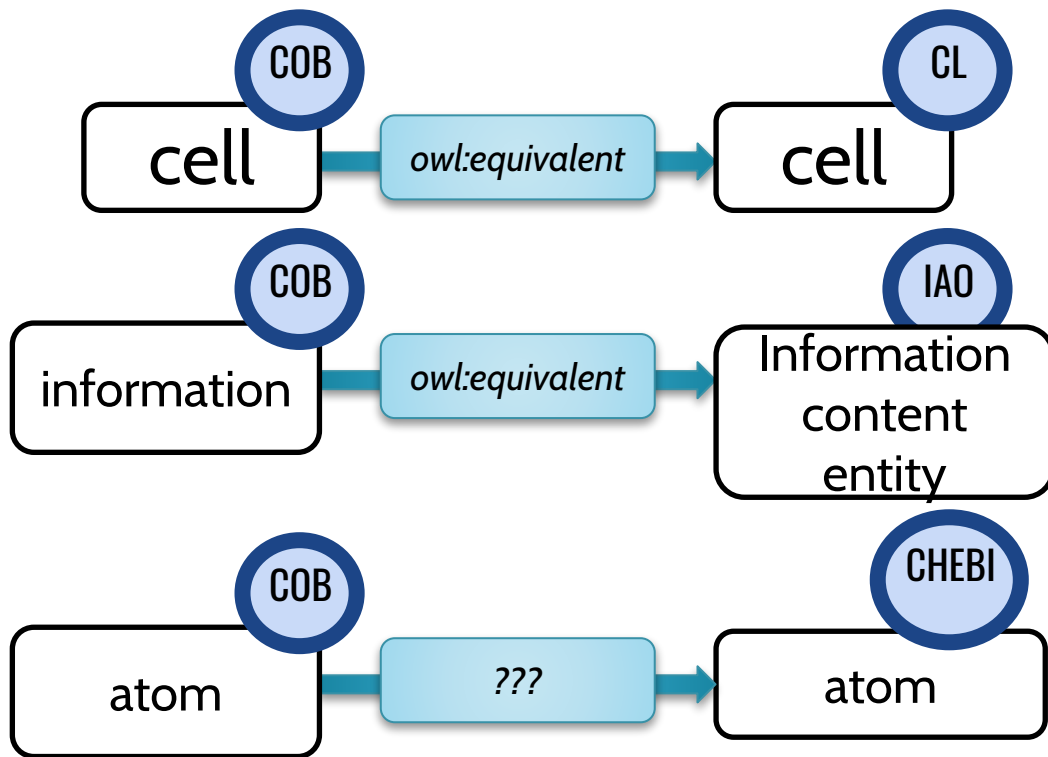
https://obofoundry.org/COB/obo-bridge

# COB and OBO top levels

- COB classes shadow existing OBO top levels

- Semantics may vary
  - If semantics are identical, we retain equivalence mappings

https://obofoundry.org/COB/obo-bridge

**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
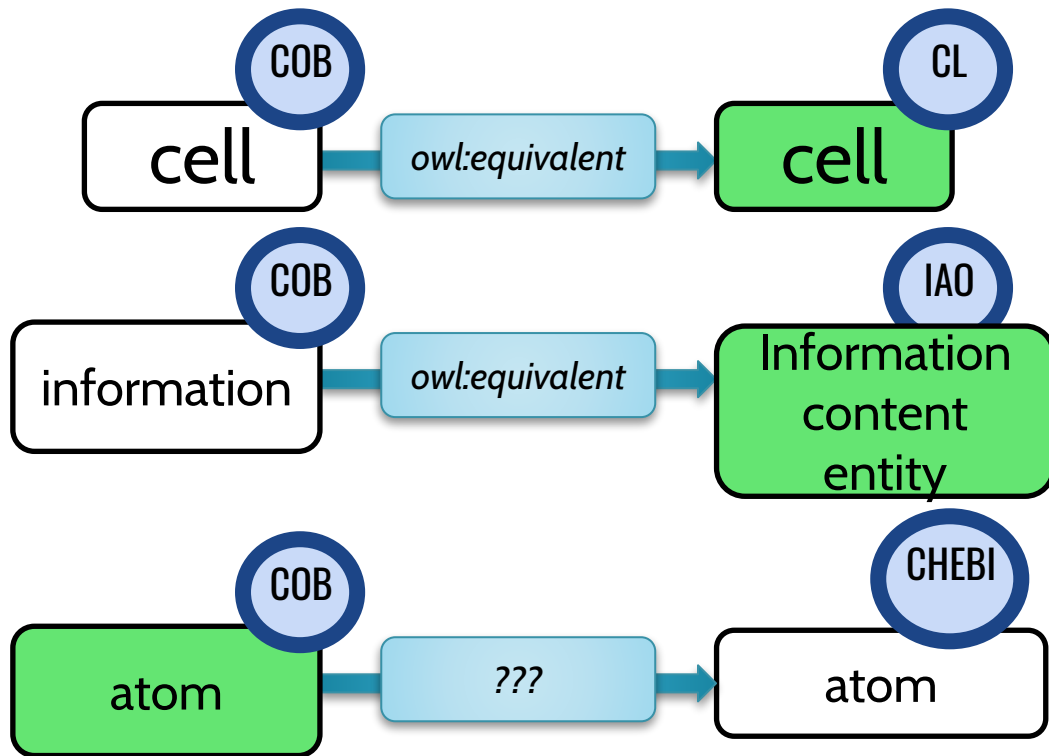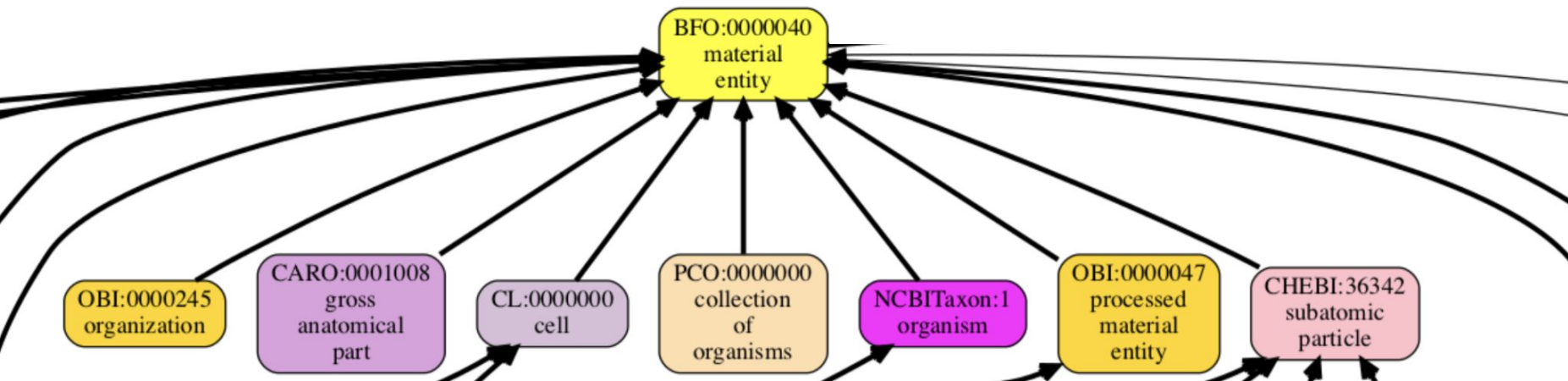
# COB development process

- 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 `OBOFoundry:master` from `cthoyt:align-chebi-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
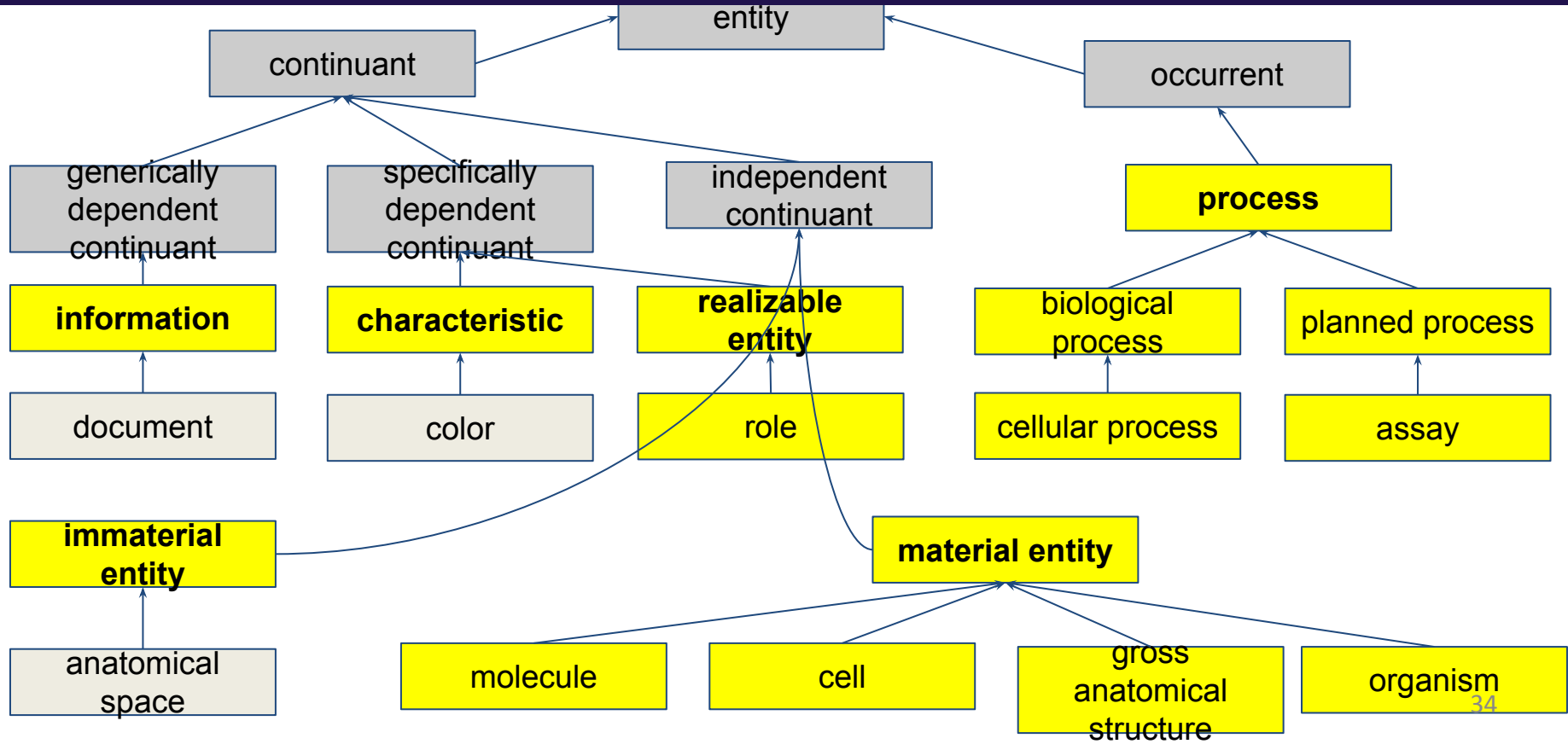
# Flattened view of BFO for COB

| process | material entity | site | characteristic | 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



smaller

*examples*

| | |
|---|---|
| subatomic particle | electron |
| atom or ion | nickel |
| molecular entity | aspirin |
| macromolecular entity | actin |
| cellular component | cell nucleus |
| cell | macrophage |
| gross anatomical structure | lung |
| organism | human |
| population | biocurators |

larger
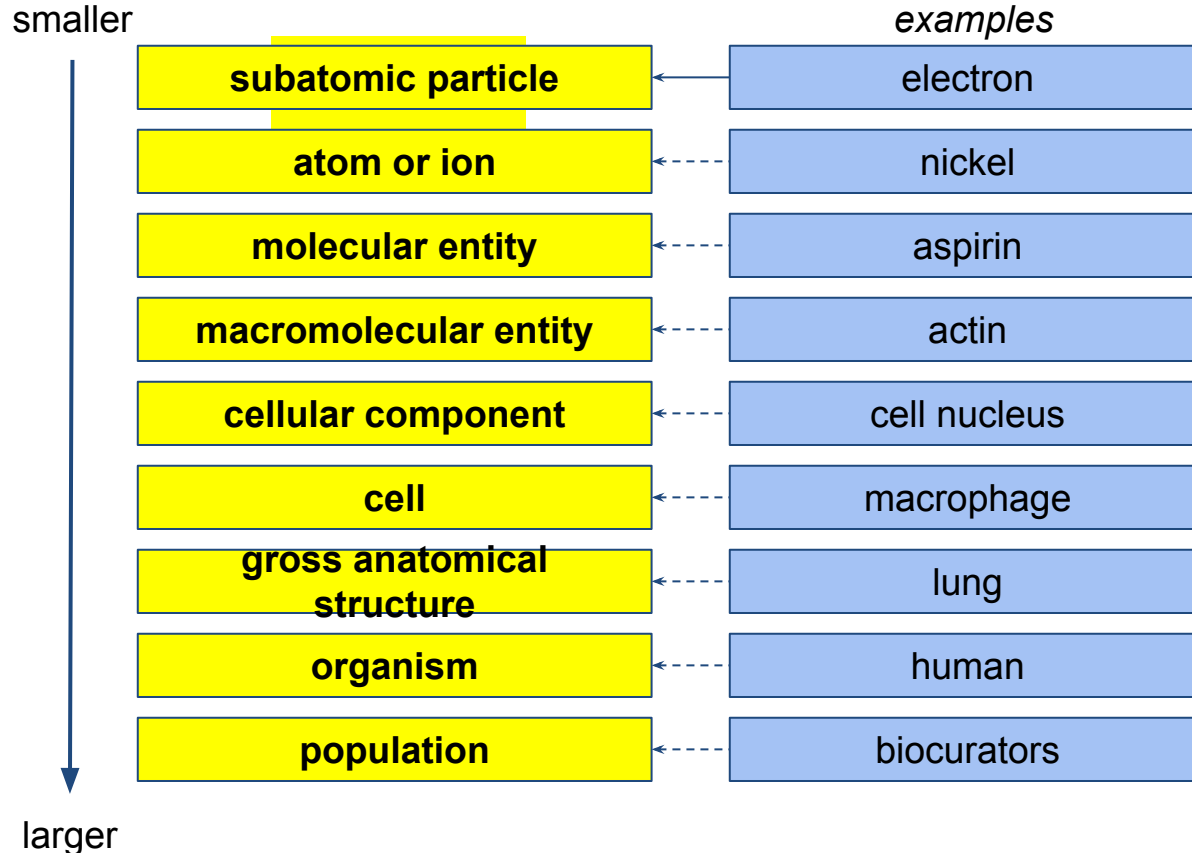
**Generic:** material entity

**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)

36

# COB – Non-Material Entities

**process** → biological process, planned process

**site** → anatomical space, geographic location

**quality** → mass, color

**characteristic** → role, plan

**information** → document, measurement datum
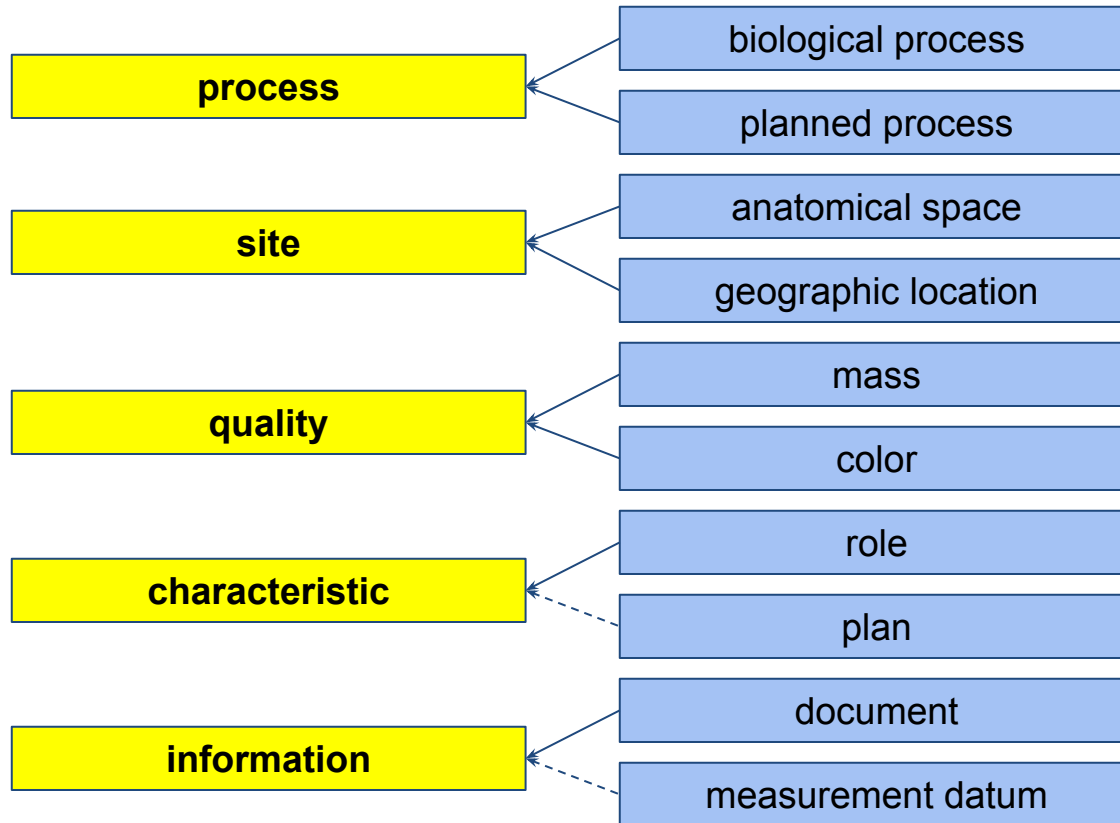
**Generic:** process, quality, realizable

**Physical world:** site, geographic location, mass, color

**Biological:** biological process, anatomical space, role

**Human activities:** planned process, information

37

# COB – Examples

## "nucleus"

| | | | |
|---|---|---|---|
| | subatomic particle | nuclear particle | **atomic nucleus** |
| material entity | cellular component | organelle | **cell nucleus** |
| | gross anatomical structure | organ | **nucleus of brain** |

## "vector"

| | | | |
|---|---|---|---|
| | organism | **infectious disease vector** | |
| material entity | processed material | **cloning vector** | |
| process | planned process | data transformation | **support vector machine** |

38
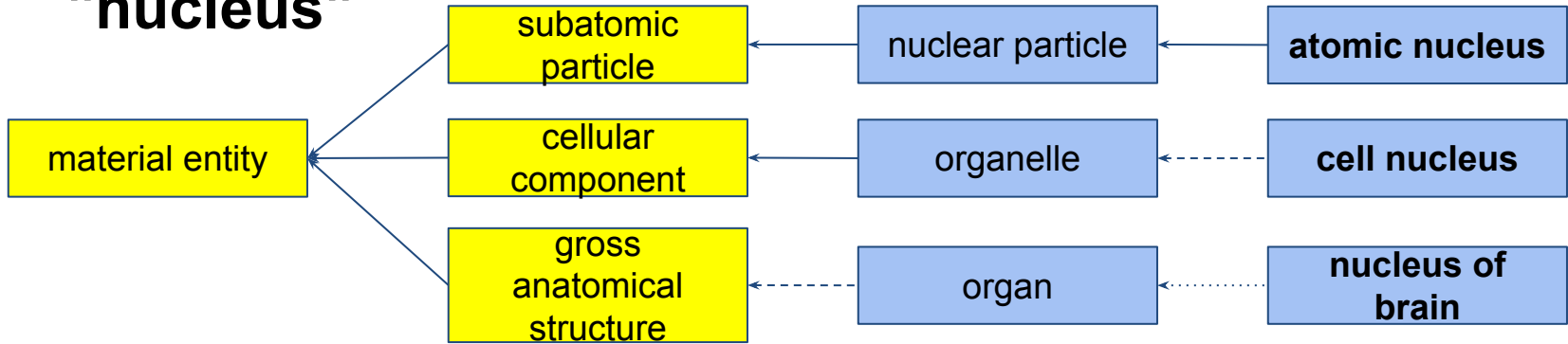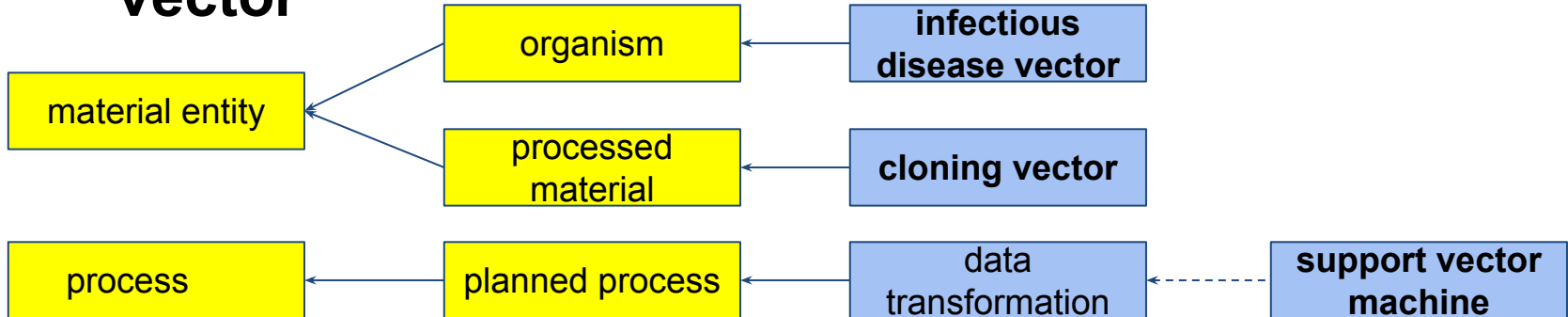
# In progress: phenotype and disease

- Upper ontology assumptions don't align with actual uses of phenotype and disease ontologies
- Consider Schulz conflation model?
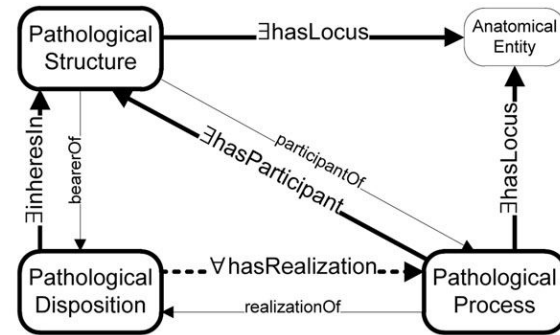
**Scalable representations of diseases in biomedical ontologies**

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

## 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
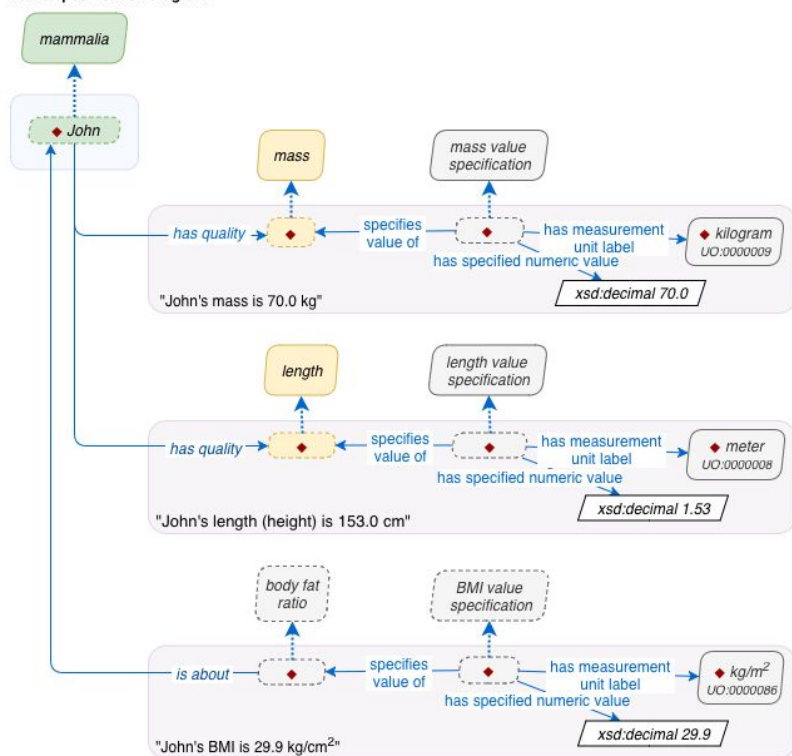
Proposed simplification of existing measurement model

# Alignment with Biolink



BioPortal

Ontologies  Search  Annotator  Recommender  Mappings

**BioLink Model**
Last uploaded: April 12, 2021

Summary  Classes  Properties  Notes  Mappings  Widgets

Jump to:

Details  Visualization  Notes ( 0 )  Class Mappings ( 37 )

- **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)*
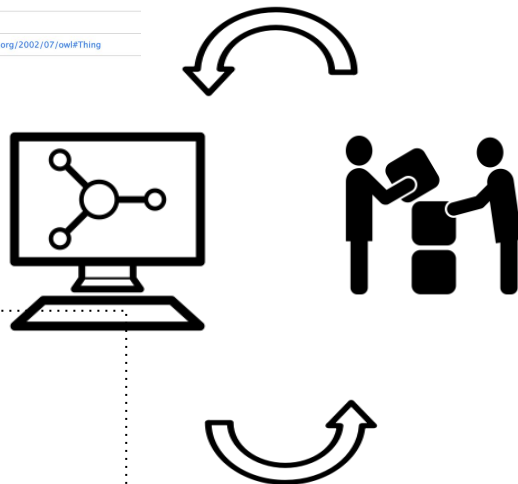
*Biolink-Model: A schema for biological KGs*
- *Expressed in LinkML*
- *"Ontology-like"*
- *OBO classes are instances*
- *Edges are first-class*

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 Matentzoglu
Anne Thessen
Damion Dooley
Chris Stoeckert

Duncan, Bill
Natale, Darren
Nomi Harris
Rebecca Jackson
Randi Vita
Yongqun He
Melissa Haendel
Barry Smith

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