

Coordinating Coronavirus Ontologies

John Beverley

johnbeve@buffalo.edu

Outline

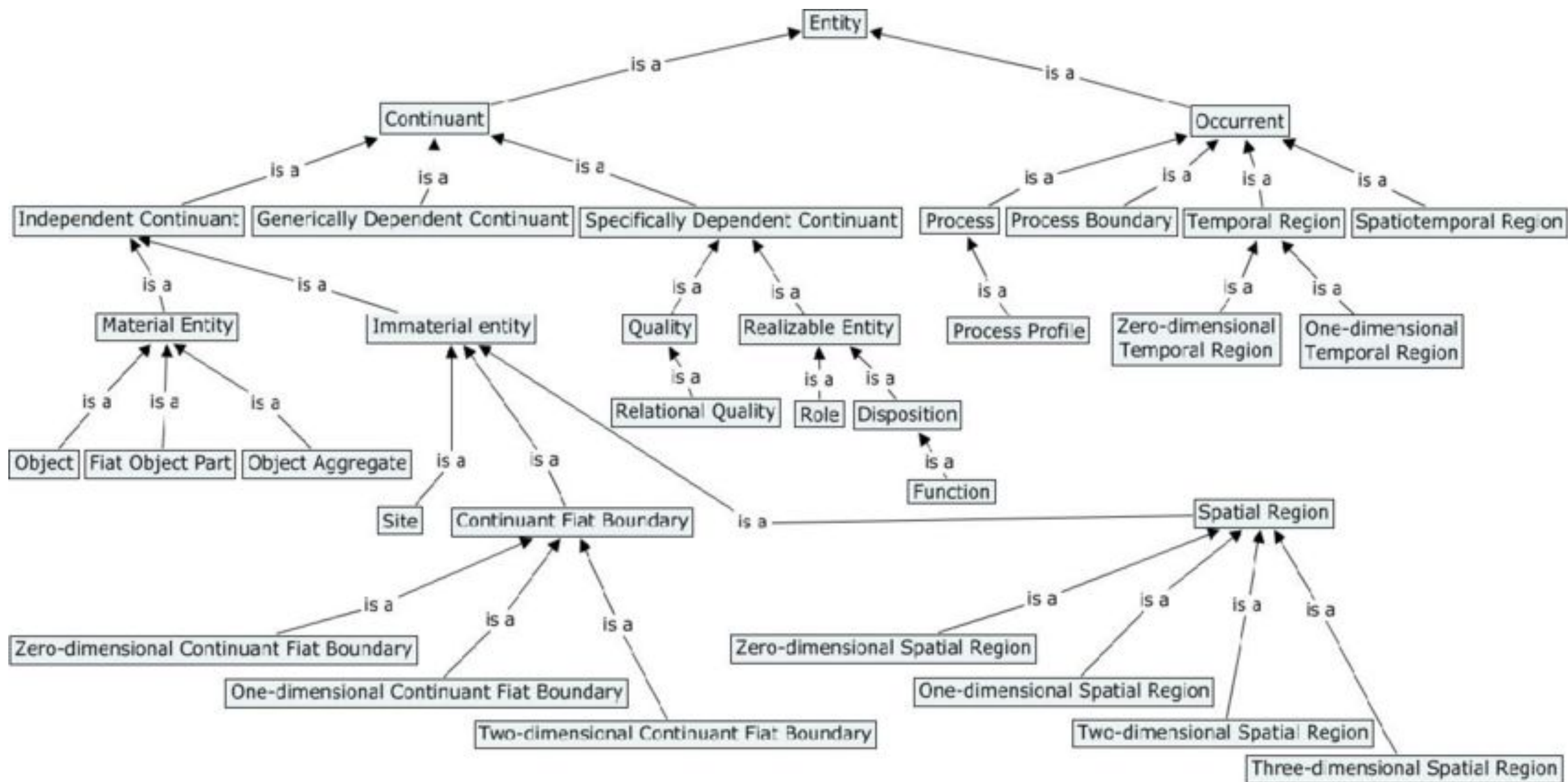
- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases
- Virus Infectious Disease Ontology
- Coordinating Coronavirus Ontologies

Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases
- Virus Infectious Disease Ontology
- Coordinating Coronavirus Ontologies

Basic Formal Ontology

- The OBO Foundry library consists of over 250 ontologies which seek to abide by the Foundry principles
- At its core is *Basic Formal Ontology*, an ISO/IEC approved standard 21838-2
- BFO is a top-level ontology covering general classes such as *material entity*, *quality*, *process*, *function*, and *role*, and provides the architecture adverted in the last Foundry principle.



Extensions of BFO

- BFO is domain-neutral, but most OBO Foundry ontologies are domain-specific
- Domain-specific ontologies represent more familiar types of entities, e.g. *disease, cell division, surgical procedure*
- Each domain ontology is constructed by downward population from BFO

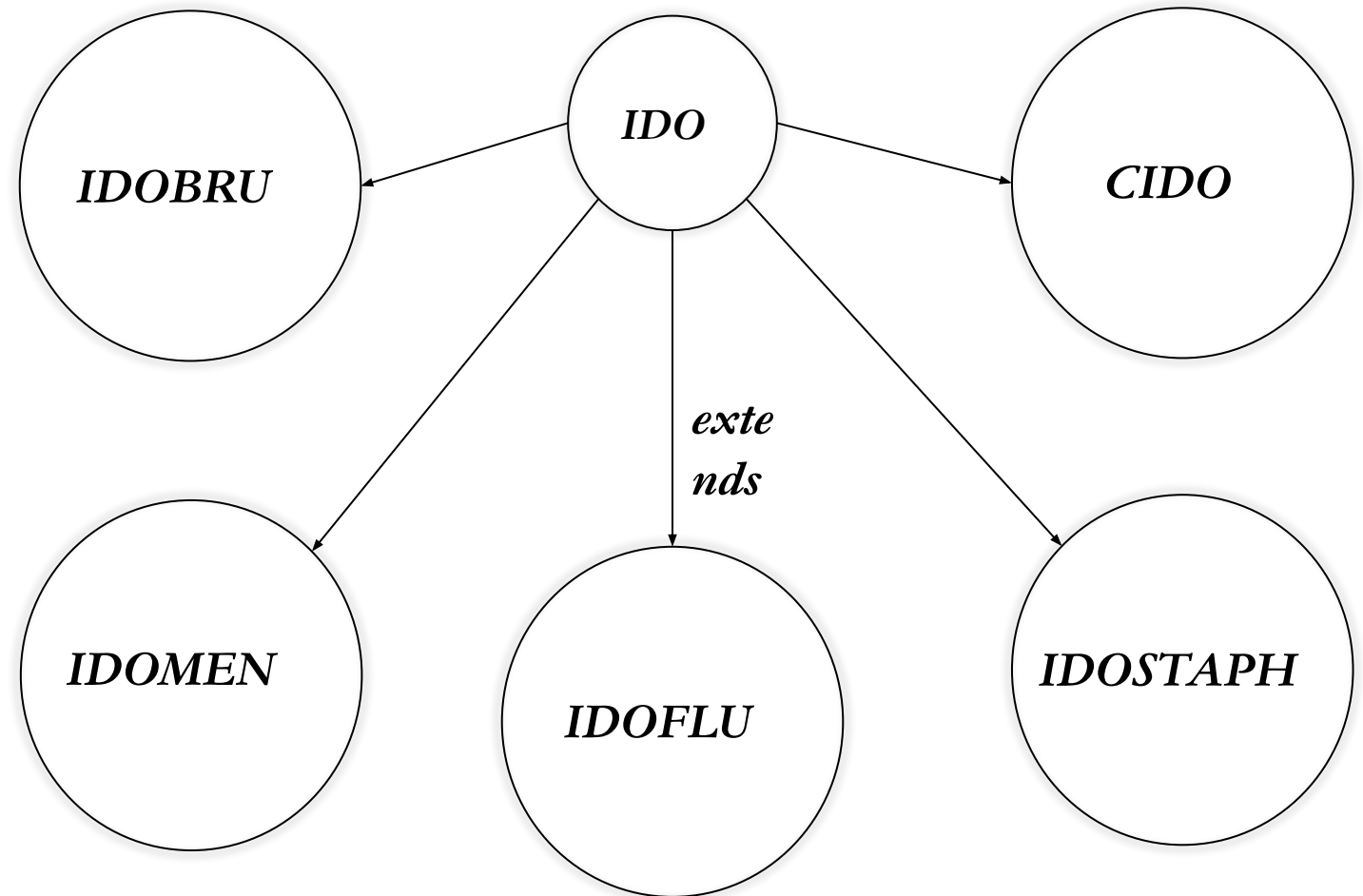
Infectious Disease Ontology

- For example, the *Infectious Disease Ontology* (IDO) *Core* is an extension of BFO to the domain of infectious diseases
- Developed by Lindsay Cowell, Albert Goldfain, and Barry Smith
- IDO Core provides researchers terms for specific infectious disease domains, e.g. *pathogen*, *asymptomatic carrier*

<https://github.com/infectious-disease-ontology>

Infectious Disease Ontology

IDO Core itself provides a starting point for ontology extensions into more specific domains



BFO Updates

- BFO underwent significant changes during ISO validation
- Consequently, IDO Core and its extensions required substantial updates
- Details of IDO Core updates can be found in a paper by Shane Babcock, Lindsay Cowell, Barry Smith, and myself titled *The Infectious Disease Ontology in the Age of COVID-19*:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8286442/>

Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases
- Virus Infectious Disease Ontology
- Coordinating Coronavirus Ontologies



Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases
- Virus Infectious Disease Ontology



<http://biportal.bioontology.org/ontologies/VIDO>

Virus Infectious Disease Ontology

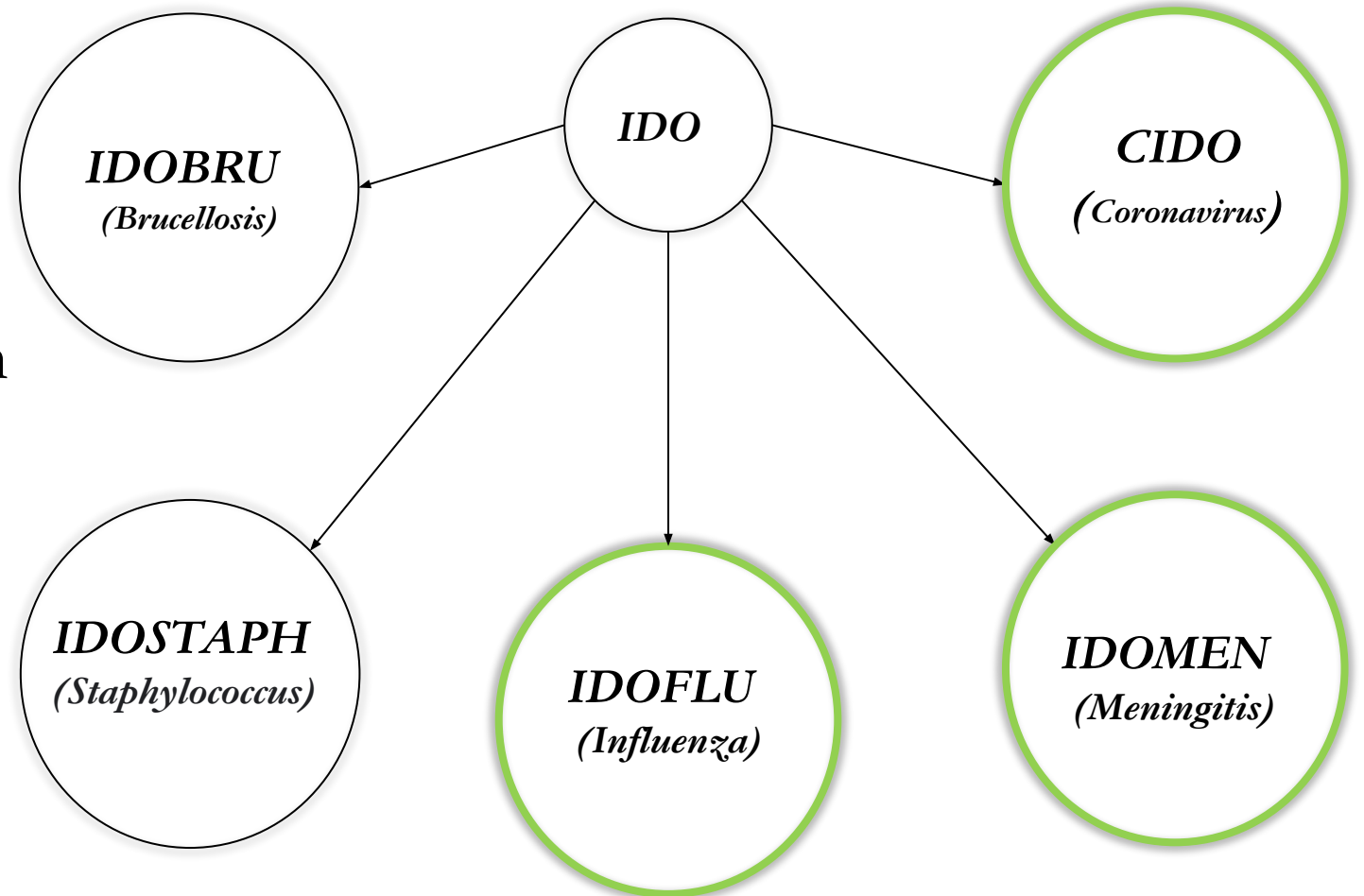
- VIDO is a reference ontology which provides a well-designed bridge between IDO Core and virus-specific extension ontologies

Details	
Acronym	VIDO
Visibility	Public
Description	The Virus Infectious Disease Ontology (IDO Virus) is an extension of the Infectious Disease Ontology (IDO). IDO Virus follows OBO Foundry guidelines, employs the Basic Formal Ontology as its starting point, and covers epidemiology, classification, pathogenesis, and treatment of terms used by Virologists, i.e. virus, prion, satellite, viroid, etc.
Status	Beta
Format	OWL
Contact	John Beverley, johnbeverley2021@u.northwestern.edu

Metrics ?	
Classes	429
Individuals	23
Properties	43
Maximum depth	9
Maximum number of children	18
Average number of children	3

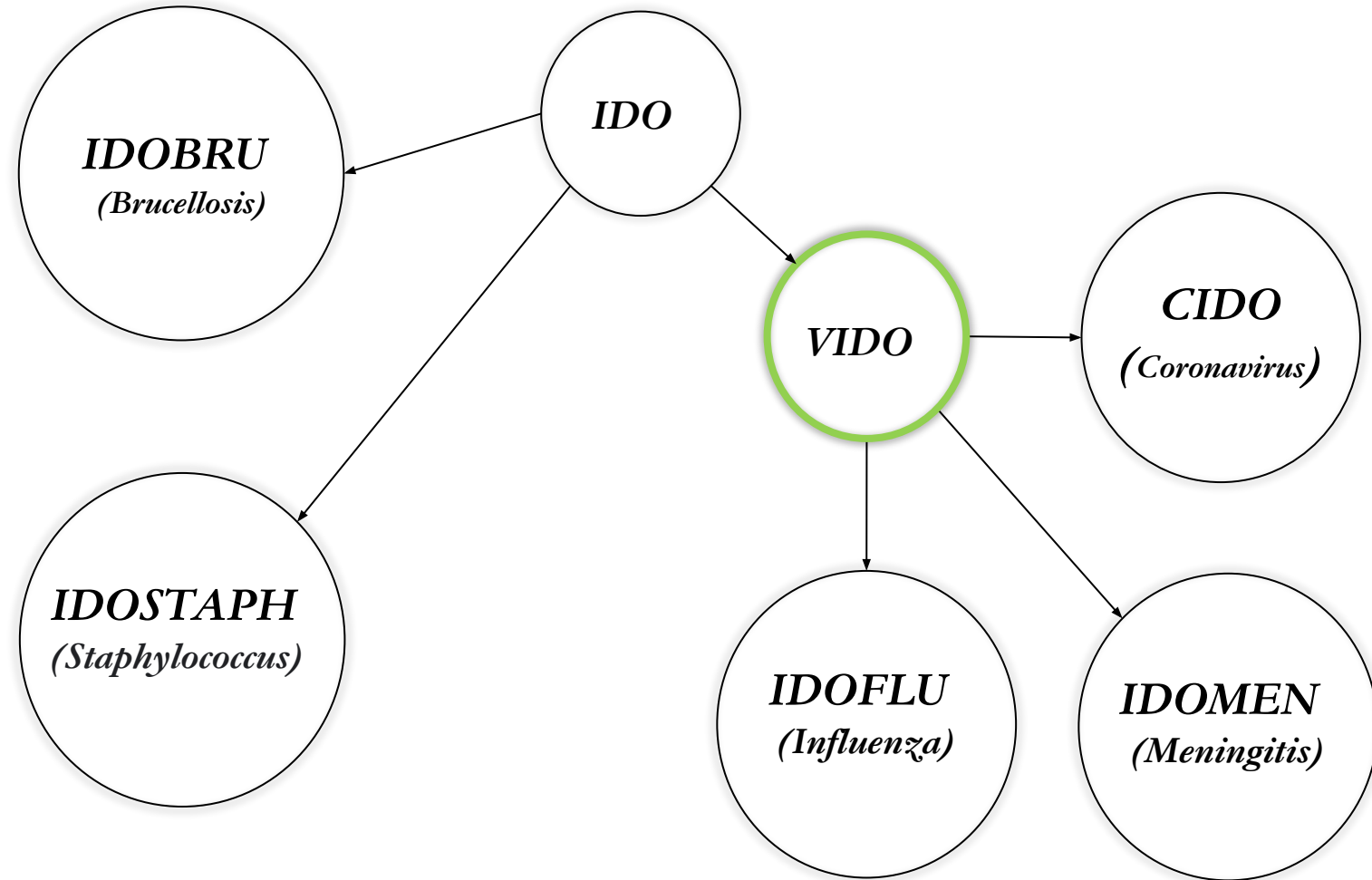
Infectious Disease Ontology

Whereas previously **virus ontologies** extended directly from IDO, and in many cases recreated ontology terms used/needed in other virus ontologies...



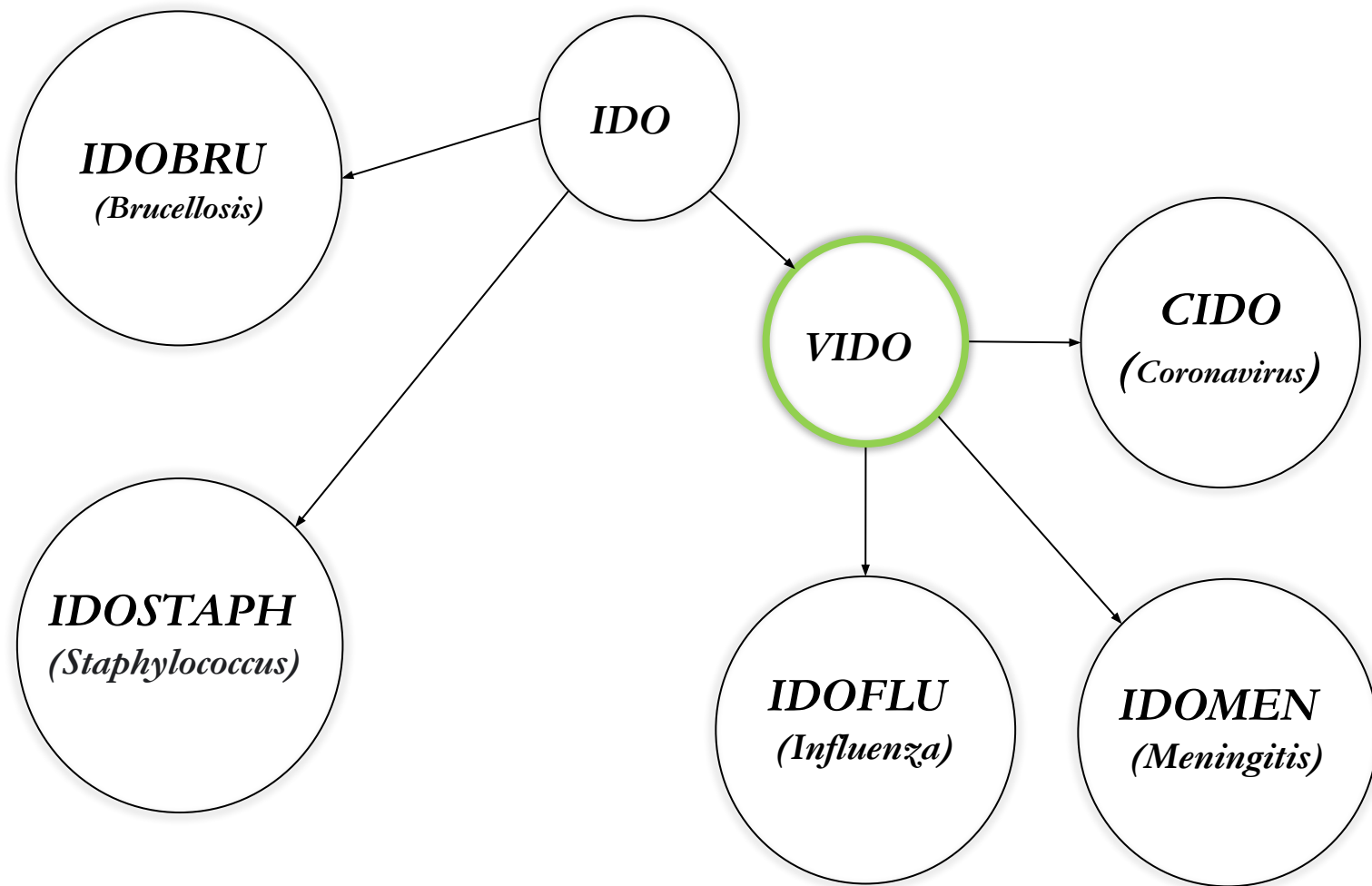
Virus Infectious Disease Ontology

...VIDO extends IDO Core
provides a bridge consisting
of terms common to most or
all virus ontologies...



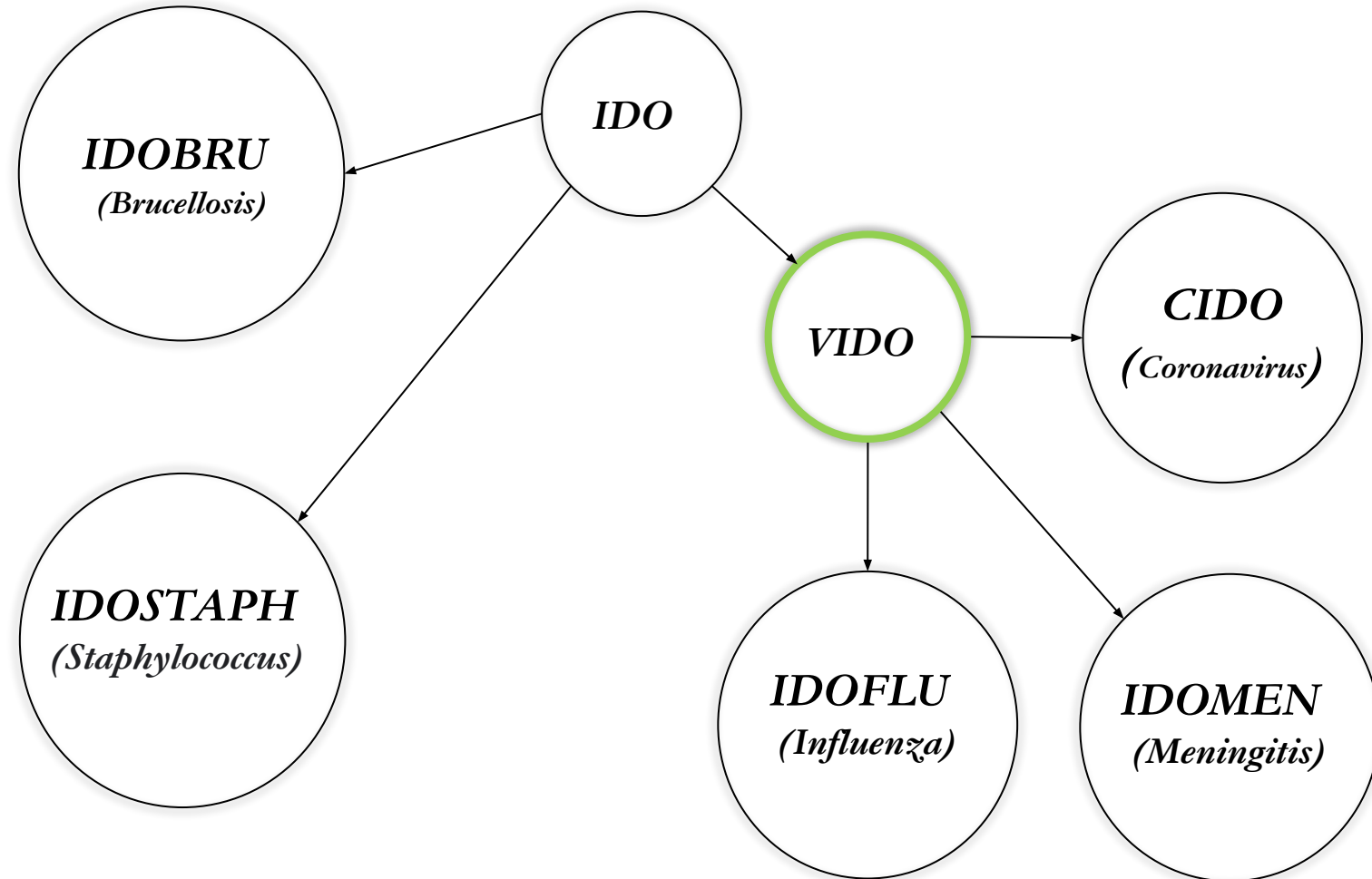
Virus Infectious Disease Ontology

...and because it was carefully designed to align with OBO principles, ensures alignment populates downward into more of the domain of virus infectious diseases.



Virus Infectious Disease Ontology

There were, of course,
ontological wrinkles to iron
out in the downward
population process...



Organisms as Cellular

- IDO Core originally counted viruses as instances of:
 - *organism* =def Object that is an individual living system, such as animal, plant, bacteria, or virus, that is capable of replicating or reproducing, growth and maintenance in the right environment. An organism may be unicellular or made up, like humans, of many billions of cells divided into specialized tissues and organs
- Where *object* - roughly, an entity that has matter as parts which exhibits causal unity – is imported from BFO

Organisms as Cellular

- IDO Core originally counted viruses as instances of:
 - *organism* =def Object that is an individual living system, such as animal, plant, bacteria, or **virus**, that is capable of replicating or reproducing, growth and maintenance in the right environment. An **organism may be unicellular** or made up, like humans, of many billions of **cells** divided into specialized tissues and organs
- Which implies viruses are cellular entities

Organisms as Cellular

- IDO Core originally counted viruses as instances of:
 - *organism* =def Object that is an individual living system, such as animal, plant, bacteria, or **virus**, that is capable of replicating or reproducing, growth and maintenance in the right environment. An **organism may be unicellular** or made up, like humans, of many billions of **cells** divided into specialized tissues and organs
- **Viruses are *acellular***

Viruses as Acellular

- Some suggested replacing *organism* with:

organism or virus or viroid

- From the Common Anatomy Reference Ontology (CARO), which is included in the OBO library
- However, this disjunctive class lumps viruses and viroids in with paradigmatic *living* entities, i.e. organisms...

Meaning of Life

- Which leads naturally to fascinating questions like:

Meaning of Life

- Which leads naturally to fascinating questions like:

Are viruses or viroids alive?

What does “life” mean?

What separates living from non-living entities?

Meaning of Life

- Which leads naturally to fascinating questions like:

Are viruses or viroids alive?

What does “life” mean?

What separates living from non-living entities?

- ...none of which obviously need answering for the purposes of *ontology modeling*

Viruses Reimagined

- IDO Core no longer counts viruses as instances of *organism*
- And instead introduces a sibling class to *organism* called *acellular structure*:

Object consisting of interrelated material parts which form an acellular unit that is the bearer of a disposition to replicate using host resources

- Under which one finds instances of virus, viroid, satellites, prions

Viruses Reimagined

- IDO Core no longer counts viruses as instances of *organism*
- And instead introduces a sibling class to *organism* called *acellular structure*:

Object consisting of interrelated material parts which form an acellular unit that is the bearer of a disposition to replicate using host resources

- Under which one finds instances of virus, viroid, **satellites**, **prions**

Virus

- VIDO imports needed terms for *virus*, *virion*, etc. from the NCBI Taxon, an OBO Foundry ontology automatically generated from the widely-used National Center for Bio**technology** database
- The NCBITaxon consists of an *exhaustive* list of biological entities
- Unfortunately...

NCBI Taxon

- NCBI Taxon does not provide:

A user-friendly hierarchy

Textual definitions

Logical definitions

- For terminological content...

NCBI Taxon

- NCBI Taxon does not provide:

A user-friendly hierarchy

Textual definitions

Logical definitions

- Its structure leads to messy ontology hierarchies

Ontofox

- Several IDO Core extension ontology developers import terms automatically using *Ontofox*
- The result is long collections of subclass relations ranging from Kingdom to Species
- Making the ontology tough to navigate

1. Data input using web forms:
Examples: [Example 1](#), [example 2](#), [example 3](#), [example 4](#), [example 5](#)

(1) Select one ontology:
Please select an ontology
Or enter your favorite source ontology and SPARQL endpoint: [Example](#)

(2) Term specification:

(a) Include low level source term URIs:
(One URI per line. To include all child terms of a source term (extract the whole branch), enter "includeAllChildren" in the line next to the source term)
Search a term:

(b) Include top level source term URIs and target direct superclass URIs (One URI per line, optional):
Search a term:

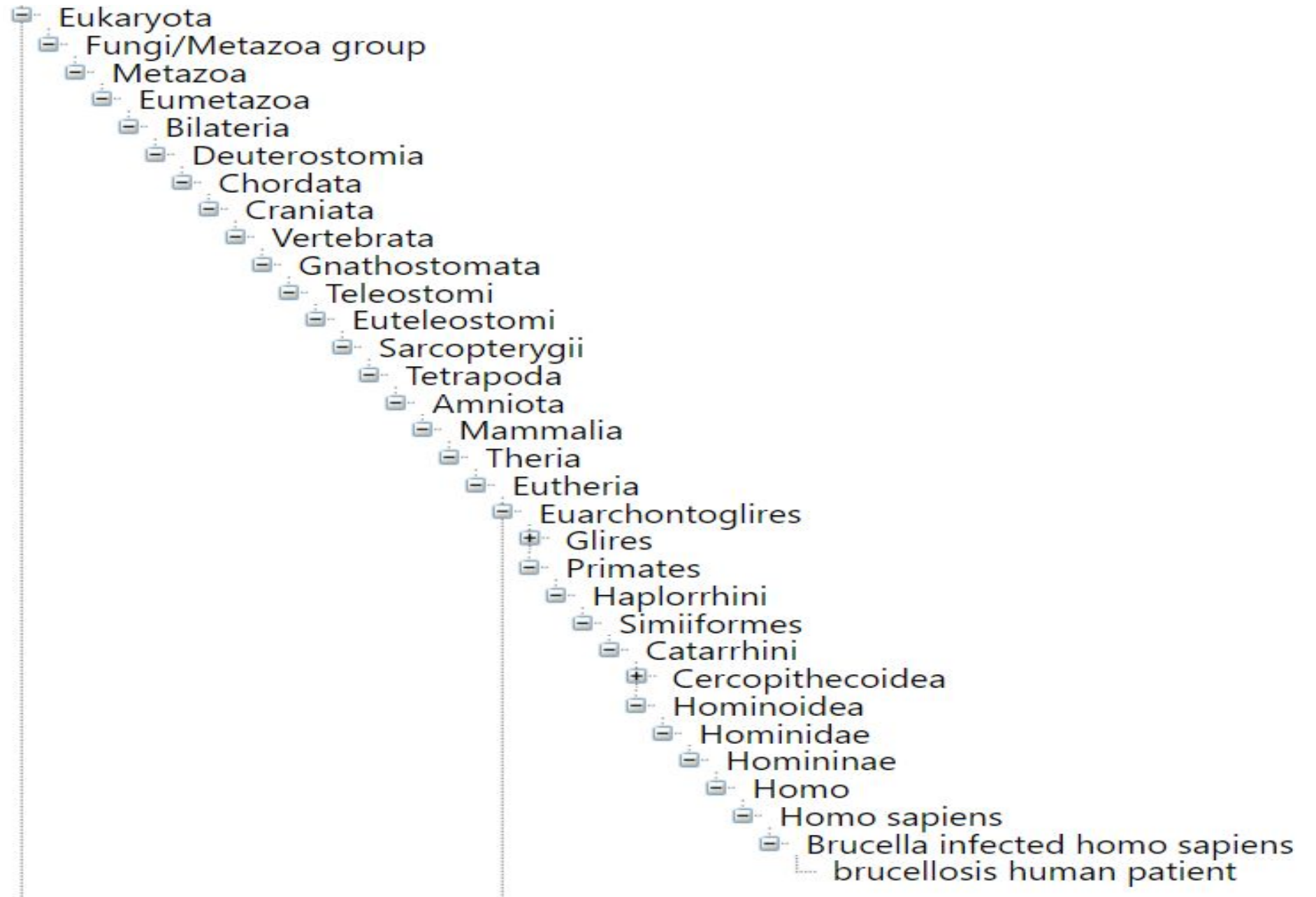
(c) Select a setting for retrieving intermediate source terms:

(3) Annotation/Axiom Specification: Include source annotation URIs (One URI per line, optional):
Examples: [rdfs:label](#), [iao:preferredTerm](#), [iao:definition](#), [iao:alternativeTerm](#), [oboInOwl:hasDefinition](#), [oboInOwl:hasSynonym](#), [owl:equivalentClass](#).
The default is no annotation to be assigned. Use [includeAllAnnotationProperties](#) to include all annotations. Use [includeAllAxioms](#) to include all annotations and other related axioms. Use [includeAllAxiomsRecursively](#) to include all axioms for the specified terms and the related terms recursively.

(4) Annotation/Axiom to be excluded (One URI per line, optional):

(5) URI of the OWL(RDF/XML) output file:
Example: http://purl.obolibrary.org/obo/vo/external/NCBITaxon_import.owl

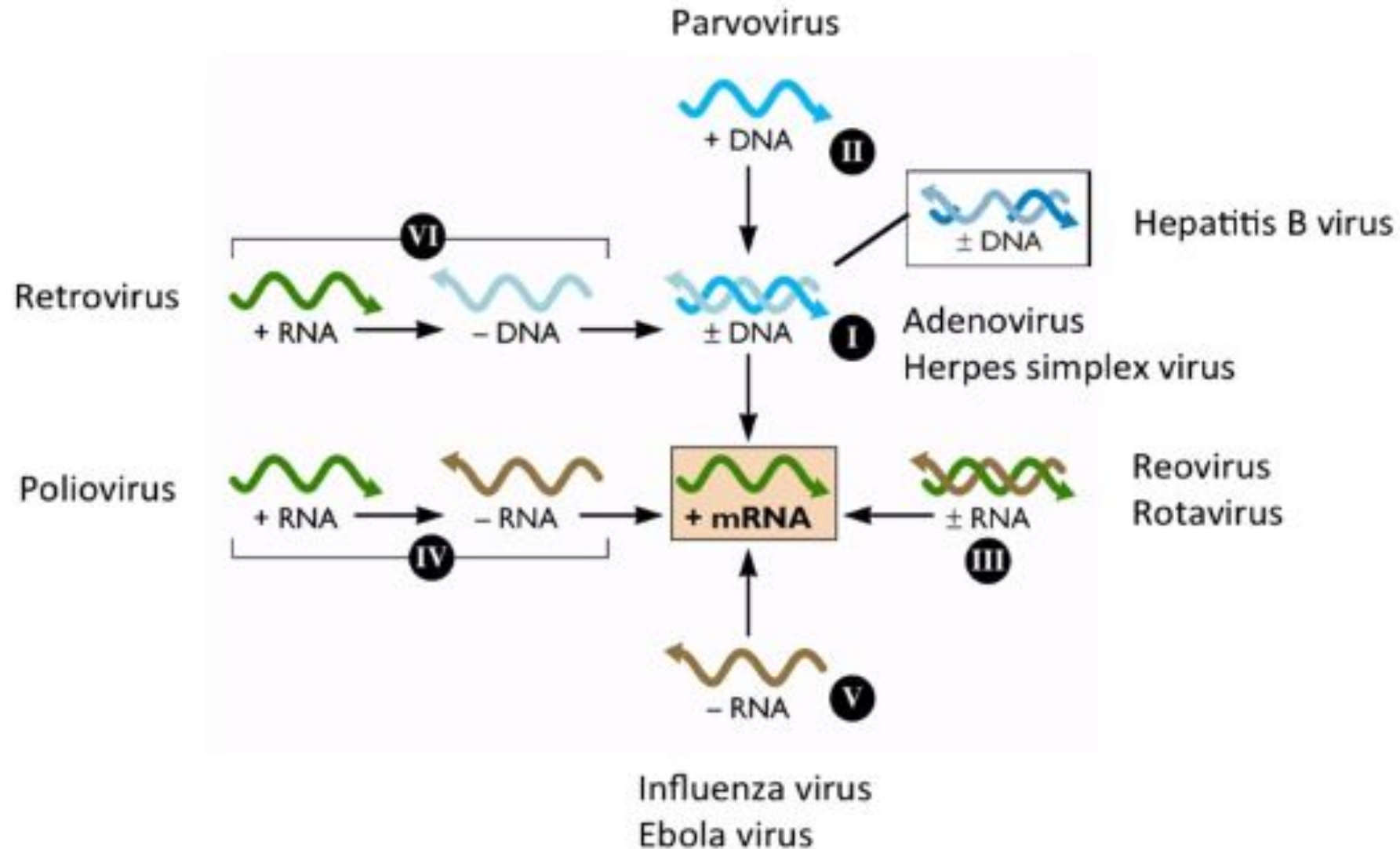
From IDOBRU, an extension of IDO Core covering brucellosis



Baltimore Classification

- Rather than a Linnean taxonomy, VIDO uses the – much simpler – *Baltimore Classification* of viruses as its starting point
 - Group I: Double-stranded DNA viruses
 - Group II: Single-stranded DNA viruses
 - Group III: Double-stranded RNA viruses
 - Group IV: Positive-sense single-stranded RNA viruses
 - Group V: Negative-sense single-stranded RNA viruses
 - Group VI: Single-stranded RNA retroviruses
 - Group VII: Double-stranded DNA retroviruses

Baltimore Classification



NCBI Taxon

- And because NCBI Taxon does not provide:

A robust hierarchical structure

Textual definitions

Logical definitions

- Re OBO principles, textual definitions had to be developed for VIDO terminological content

Virus

- *virus* =def Acellular structure with RNA or DNA genetic material which relies on host metabolic resources for RNA or DNA replication
- Alongside groups in the Baltimore Classification, such as:
- *positive-sense single-stranded RNA virus* =def Virus with genetic material encoded in positive-sense single-stranded RNA that can be translated directly into proteins



- infectious agent reservoir
- macromolecular complex
- molecular entity
- object
 - acellular structure
 - disordered prion
 - disordered virus
 - infectious structure
 - prion
 - virus
 - bacteriophage
 - double-stranded DNA retrovirus
 - double-stranded DNA virus
 - double-stranded RNA virus
 - mycovirus
 - negative-sense single-stranded RNA virus
 - oncovirus
 - **positive-sense single-stranded RNA virus**
 - single-stranded DNA virus
 - single-stranded RNA retrovirus
 - virion

● positive-sense single-stranded RNA virus — NCBITaxon:35278 — <http://purl.obolibrary.org/ok>

Annotations: positive-sense single-stranded RNA virus

Annotations +

label [language: en]

positive-sense single-stranded RNA virus

definition

Virus with genetic material encoded in single-stranded RNA that can be translated directly into proteins.

hasExactSynonym

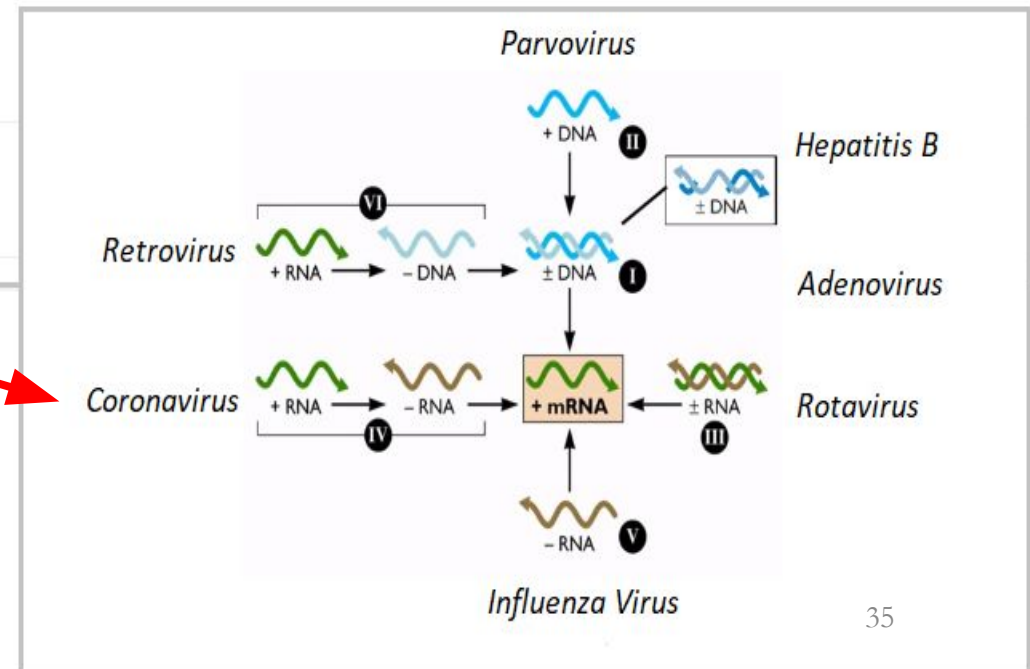
Baltimore Classification Group IV

'example of usage'

Coronavirus

'example of usage'

Picornavirus





- infectious agent reservoir
- macromolecular complex
- molecular entity
- object
 - acellular structure
 - disordered prion
 - disordered virus
 - infectious structure
 - prion
 - virus
 - bacteriophage
 - double-stranded DNA retrovirus
 - double-stranded DNA virus
 - double-stranded RNA virus
 - mycovirus
 - negative-sense single-stranded RNA virus
 - oncovirus
 - **positive-sense single-stranded RNA virus**
 - single-stranded DNA virus
 - single-stranded RNA retrovirus
 - **virion**

● positive-sense single-stranded RNA virus — NCBITaxon:35278 — <http://purl.obolibrary.org/ok>

Annotations: positive-sense single-stranded RNA virus

Annotations +

label [language: en]

positive-sense single-stranded RNA virus

definition

Virus with genetic material encoded in single-stranded RNA that can be translated directly into proteins.

hasExactSynonym

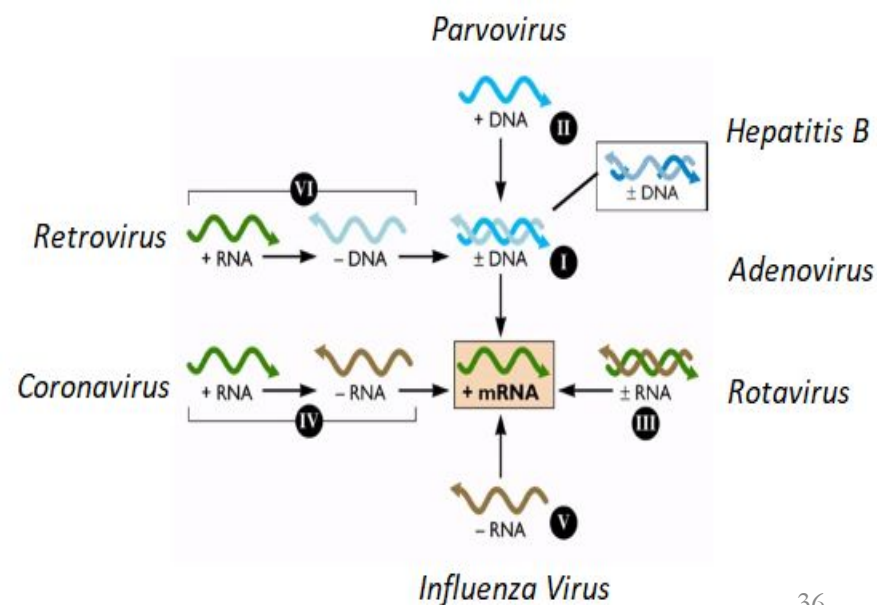
Baltimore Classification Group IV

'example of usage'

Coronavirus

'example of usage'

Picornavirus



Virion

- *virion* =def Virus in its assembled state consisting of genomic material surrounded by coating molecules

Virion

- *virion* =def Virus in its assembled state consisting of genomic material surrounded by coating molecules
- Options for understanding relationship between virion and virus:

Virion is to virus as virus is to virion (synonyms)

Virion is to virus as (human) gamete is to human

Virion is to virus as (human) student is to human

Virion is to virus as (human) child is to human

Virion

- *virion* =def Virus in its assembled state consisting of genomic material surrounded by coating molecules
- Options for understanding relationship between virion and virus:

Virion is to virus as virus is to virion (synonyms)



Virion is to virus as (human) gamete is to human



Virion is to virus as (human) student is to human



Virion is to virus as (human) child is to human



Viruses as Pathogens

- Viruses are often described as being “obligate pathogens”
- The term “pathogen” is not used consistently in life science literature
- My proposal to unify the ambiguity is to first understand “pathogen” as indexed either to a species or to stages in the developmental cycle of a species
- And define a pathogen as an entity bearing a certain *disposition*

Pathogenic Disposition

pathogenic disposition =def Disposition borne by a material entity to establish localization in or produce toxins that can be transmitted to, an organism, either of which may form disorder in the organism or immunocompetent members of the organism's species.

Pathogenic Disposition

pathogenic disposition =def Disposition borne by a material entity to establish **localization in** or produce toxins that can be transmitted to, an organism, either of which may form disorder in the organism or immunocompetent members of the organism's species.

SARS-CoV-2 attachment
to host cell; S. Aureus
opportunistic pathogens

Pathogenic Disposition

Host ingests toxin
from *C. botulinum*



pathogenic disposition = def Disposition borne by a material entity to establish **localization in** or **produce toxins that can be transmitted to**, an organism, either of which may form disorder in the organism or immunocompetent members of the organism's species.



SARS-CoV-2 attachment
to host cell; *S. Aureus*
opportunistic pathogens

Pathogenic Disposition

Host ingests toxin
from *C. botulinum*



pathogenic disposition = def Disposition borne by a material entity to establish **localization in** or **produce toxins that can be transmitted to**, an organism, either of which may form disorder in the organism or **immunocompetent members of the organism's species**.



SARS-CoV-2 attachment
to host cell; *S. Aureus*
opportunistic pathogens



HIV-1 infection in
host with CCR-5 mutation;
p. falciparum infection in
host with sickle-cell trait

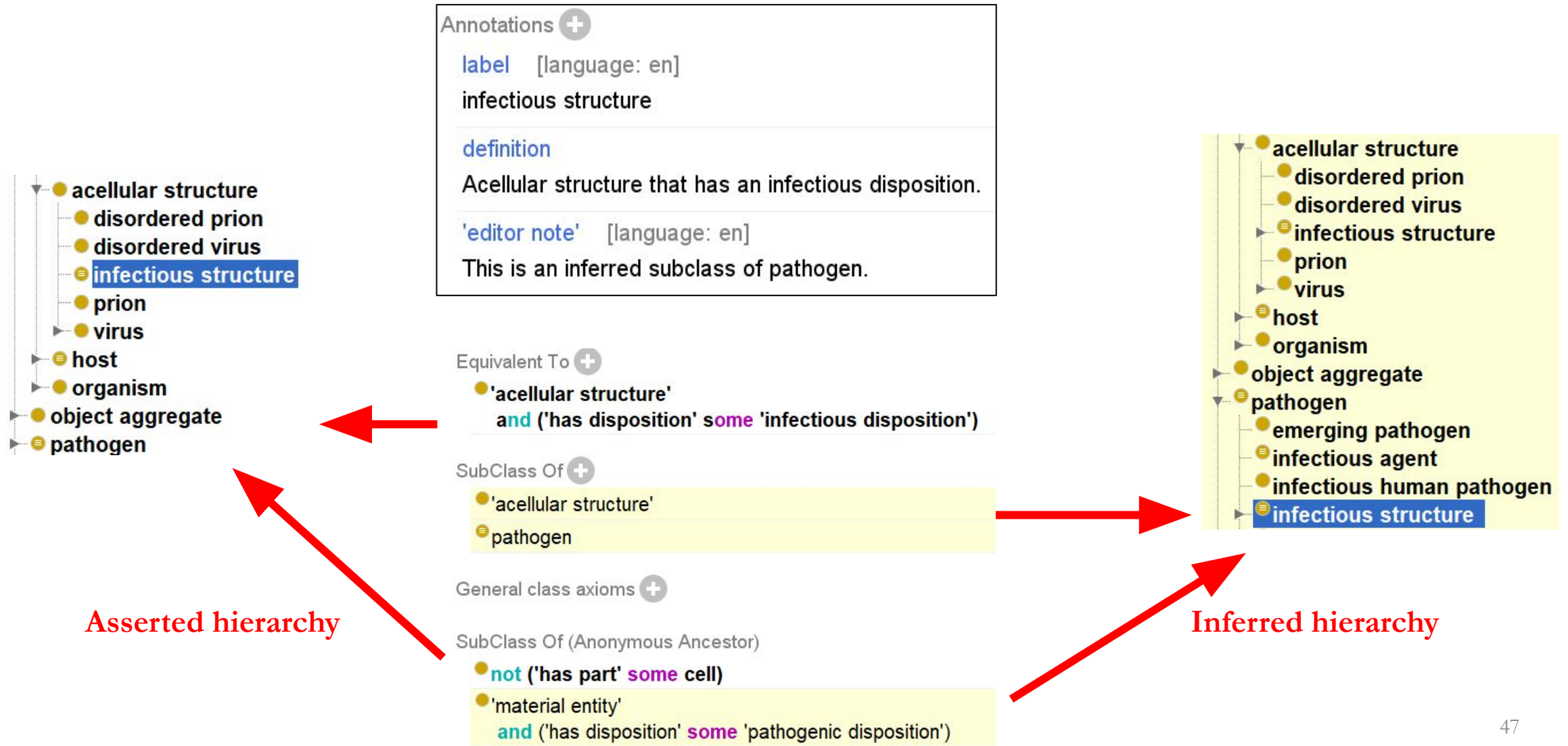
Infectious Disposition

- Viruses are not merely pathogens, they're often infectious pathogens
- *infectious disposition* =def Pathogenic disposition borne by a pathogen to be transmitted to a host and then become part of an infection in that host or immunocompetent members of the same species as the host.
- SARS-CoV-2, for example, not only localizes and causes disorder in a host, but is disposed to transmit to and become part of an infection in a host

Infectious Structures

- Any acellular structure – like a virus – bearing an infectious disposition, counts as an *infectious structure*
- Because infectious dispositions are a subclass of pathogenic dispositions, any infectious structure is a pathogen
- However, not all viruses are infectious structures, e.g. disordered viruses may not be infectious
- We use the inferred hierarchy to reflect these distinctions

Infectious Structures



Viral Pathogenesis

- A virus realizing an infectious disposition involves ordered processes:

Appearance of disorder

Process of establishing infection

Establishment of localization in host

Transmission process

Viral Pathogenesis

- Axiom assertion capabilities are a common feature of ontology tools:

Appearance of disorder

Process of establishing infection

Establishment of localization in host

Transmission process

Asserted

Inferred

SubClass Of +

'part of' some pathogenesis

'preceded by' some 'establishment of localization in host'

process



General class axioms +

SubClass Of (Anonymous Ancestor)

'part of' some pathogenesis

'preceded by' some 'pathogen transmission process'

Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases 
- Virus Infectious Disease Ontology 
- Coordinating Coronavirus Ontologies

Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases 
- Virus Infectious Disease Ontology 
- Coordinating Coronavirus Ontologies

Ontology Harmonization

- VIDO was developed alongside a BFO-conformant ontology designed to extend IDO to coronaviruses
- COVID-19 Infectious Disease Ontology (IDO-COVID-19)

<https://biportal.bioontology.org/ontologies/IDO-COVID-19>

Ontology Harmonization

- VIDO was developed alongside a BFO-conformant ontology designed to extend IDO to coronaviruses
- COVID-19 Infectious Disease Ontology (IDO-COVID-19)

<https://biportal.bioontology.org/ontologies/IDO-COVID-19>

- IDO-COVID-19, however, overlapped existing coronavirus ontologies

Ontologies in the Age of COVID-19

IDO-COVID-19
was not unique in that respect;
searching Bioportal for
“COVID-19” returns several
coronavirus ontologies,
developed without coordination,
often overlapping

COVID-19 Surveillance Ontology (COVID19)

The COVID-19 Surveillance Ontology is an application ontology used to support COVID-19 (Wuhan novel Coronavirus infection) surveillance in primary care

Uploaded: 3/22/20

projects
1

classes
52

COVID-19 Ontology (COVID-19)

The COVID-19 ontology covers the role of molecular and cellular entities in virus-host-interactions, in the virus life cycle, as well as a wide spectrum of medical and epidemiological concepts linked to COVID-19.

Uploaded: 6/22/20

classes
2,268

COVID-19OntologyInPatternMedicine (COVID-19-ONT-PM)

This ontology focuses on the aspects of the scientific findings in the COVID-19 fields in order to help the efforts against the outbreak of the pandemic

Uploaded: 11/19/20

projects
1

classes
136

WHO COVID-19 Rapid Version CRF semantic data model (COVIDCRFRAPID)

This is a semantic data model for the WHO's COVID-19 case record form RAPID version from April 8 2020

Uploaded: 6/24/20

projects
1

classes
398

The COVID-19 Infectious Disease Ontology (IDO-COVID-19)

The COVID-19 Infectious Disease Ontology (IDO-COVID-19) is an extension of the Infectious Disease Ontology (IDO) with COVID-19 related concepts.

classes
486

Coronavirus Infectious Disease Ontology

- Of particular note, is the OBO Coronavirus Infectious Disease Ontology ([CIDO](#)) – developed by Yongqun “Oliver” He – which covers coronavirus diseases, details of which can be found in the article below:

CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis

Yongqun He , Hong Yu, Edison Ong, Yang Wang, Yingtong Liu, Anthony Huffman, Hsin-hui Huang, John Beverley, Junguk Hur, Xiaolin Yang, Luonan Chen, Gilbert S. Omenn, Brian Athey & Barry Smith

Scientific Data **7**, Article number: 181 (2020) | [Cite this article](#)

4103 Accesses | **4** Citations | **17** Altmetric | [Metrics](#)



Ontology Harmonization

- Asiyah Yu Lin spearheaded a harmonization effort between developers of CIDO, IDO-COVID-19, and several other COVID-19 ontologies
- As a result, IDO-COVID-19 was subsumed by CIDO and plans for harmonizing CIDO with VIDO have been established



**Image from COVID-19 Ontology Harmonization poster ICBO 2021, Asiyah Yu Lin, et al.*

Outline

- Open Biomedical and Biological Ontologies Foundry, Basic Formal Ontology, and infectious diseases 
- Virus Infectious Disease Ontology 
- Coordinating Coronavirus Ontologies 