

Challenges Arising from Ontology Imports Utilized for Exploring Mechanisms of Disease



J. Allen Baron and Lynn M. Schriml*

University of Maryland School of Medicine, Institute for Genome Sciences, Baltimore, MD, USA

* Presenting Speaker

www.disease-ontology.org

github.com/DiseaseOntology

[@diseaseontology](https://twitter.com/diseaseontology)

ABSTRACT

The Human Disease Ontology (DO) seeks to describe the breadth and complexity of human disease and to provide a stable framework for advanced analysis. It accomplishes these goals through automated import of non-disease ontologies and the definition of logical axioms. Exploration of diseases via these imports and axioms is available by browsing DO's OWL tree, through advanced searches at DO's website, or by downloading the doid.owl file from DO's GitHub repository. Utilizing other ontologies to define disease mechanisms contributes to DO's interoperability but also presents challenges. Multiple examples of challenges faced while curating the DO and the approaches used to mitigate them are outlined here. These address common problems in ontology curation including ontology size, differing scopes/philosophies, unexpected changes, and ultimately the real-world difficulty of defining (medical) terms.

ADVANCED SEARCH

Advanced Search on disease-ontology.org enables powerful exploration of diseases, returning complete or partial matches across DO and DO's imports.

Advanced Search

✓ Name

Synonym

Definition

Subset

DOID

Alternate ID

Xrefs

Relation

i.e. class label (ex: "foot")

limited to synonyms in DO

ex: DO_cancer_slim

ID (ex: "FOODON:0", *partial match*)

cross-reference ID (ex: "OMIM:253550", *full match*)

any RO relation term in DO; *only DOID results*

Relation search example

Advanced Search

Relation

disease has location

Keyword:

finger

ID

Name

DOID:0050581 brachydactyly

DOID:0110973 Mononen-Karnes-Senac syndrome

MANAGING IMPORT CHALLENGES

Custom import of ontologies is powered by curation and the ROBOT tool. Custom "import files" (*.txt files, [turquoise](#)) specifying terms to import from other OBO Foundry ontologies are generated & maintained by the DO curators. ROBOT commands (code blocks, variable/optional components [yellow](#)) sync the imports during each DO release with their source owl files. The ROBOT commands to create a DO import are: extract - remove - annotate

CHALLENGE: Large Size of Imports

DO incorporates over a dozen imports, several of which are quite large and could slow reasoning of the doid.owl file (example: Chemical Entities of Biological Interest (ChEBI) has 166,001 classes and a file size of 612MB).

FIX: Curate a list of terms relevant to DO in a custom "import file" for import via robot extract (example: [chebi_terms.txt](#)).

```
robot extract
--input <chebi.owl>
--method mireot
(--upper-term <not used here>)
--lower-terms <chebi_terms.txt>
```

The imports and ROBOT commands are available in DO's GitHub import directory: github.com/DiseaseOntology/HumanDiseaseOntology/tree/main/src/ontology/imports

CHALLENGE: Overlapping Scope in Import

A disease can be described by the co-occurrence of 'disease features', some of which are phenotypes and others diseases. The Human Phenotype Ontology (HPO) includes diseases as phenotypes when they are a feature of another disease. Importing these phenotype terms would duplicate DO diseases (example: diabetes mellitus, DOID:9351 & HP:0000819).

FIX: Codify these unique disease relationships in DO with axioms defined using the RO (Relation Ontology) relation '**disease has feature**', while minimizing scope conflicts with HPO by importing the 'Phenotypic Abnormality' branch and programmatically via an 'exclude' file, removing them from the HPO import file.

robot extract <HPO phenotypic abnormality>

remove --term-file exclude-hp.txt

diabetes mellitus

inguinal hernia

HP:0000819

HP:0000023

...

CHALLENGE: Incompatible Import Term Labels

The cross-species anatomy ontology UBERON may use non-human default labels for anatomical parts. **Example:** UBERON:0002387 "pes" → "zoological term for the distal portion of the hind limb of tetrapod animals"

Pes

→

Foot

uberon_terms.txt

UBERON:0002387 # foot

UBERON:0002398 # hand

...

FIX: Update the labels to the human anatomical terms in the [uberon_terms.txt](#) import file.

CHALLENGE: Finding a Home for New Imports

To model complex disease, new terms that describe environmental drivers of disease are needed. These are outside the scope of DO (not diseases).

FIX: Determine the set of terms describing disease drivers and align them to Exposure Ontology (ExO) upper level terms, and add an ExO import to the DO. We requested these terms be added to ExO, which includes the parent concepts of these 100+ stressor terms, but the scope of ExO is limited to upper-level terms. As adding these terms to the related Environmental Conditions,

assay

exposure outcome

exposure stressor

biological agent

biomechanical agent

chemical agent

ecological perturbation

acid rain

climate change

cyclone

earthquake

fire

thunderstorm

tornado

tsunami

physical agent

psychosocial agent

poverty

source

exposure_receptor

process

exposure stressor

biological agent

chemical agent

ecological perturbations

acid rain

climate change

cyclone

earthquake

fire

thunderstorm

tornado

tsunami

physical agent

extreme heat

noise

radiation

strenuous exercise

ultraviolet light

psychosocial agent

ExO

Treatments and Exposures Ontology (ECTO) was also not an option (too many terms), thus the new Disease Drivers ontology was created to import needed terms.

Disease Drivers (DISDRIV)