The DO-KB is expanding the disease information ecosystem
- expanding the representation of the DO diseaseome
- a comprehensive network of disease to disease relationships
- a disease feature similarity network for differential diagnosis exploration

**DO-KB SPARQL Sandbox**
- A data playground for querying the Human Disease Ontology SPARQL endpoint and to perform federated SPARQL queries across connected resources for disease knowledge discovery.
- A linked data representation (RDF triple store) of the Human Disease Ontology's diseases, disease features and mechanisms.

**DO-KB Faceted Search Interface**
- A new way to explore human disease features and mechanisms, represented here as data facets.
- The DO-KB Faceted Search Interface enables exploration of connectivity across diseases.

The DO-KB's SPARQL endpoint, explore DO data and query disease, genomic, proteomic resources through federated queries.

Mine the DO-KB disease-data connections through faceted query and retrieval of DO diseases sharing features or mechanisms, including phenotypes, environmental or genetic drivers, anatomy, variant type, and age of onset.
**What is SPARQL?**

- **SPARQL:** SPARQL Protocol And RDF Query Language
- An RDF (Resource Description Framework) Query Language
- A semantic query language for databases

**What can you do with SPARQL?**

- Enables knowledge to be represented in a machine-readable way
- Used to retrieve and manipulate data stored in RDF format.
- Codifies (stores) relationships between semantic data, in the form of subject-predicate-object ‘triples’. E.g., Baltimore is in Maryland
- Use SPARQL to create complex queries with SELECT statements
- A SPARQL endpoint which allows you to query against the data set

<table>
<thead>
<tr>
<th>service</th>
<th>endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO</td>
<td><a href="https://sparql.proconsortium.org/virtuoso/sparql">https://sparql.proconsortium.org/virtuoso/sparql</a></td>
</tr>
<tr>
<td>UniProt</td>
<td><a href="http://sparql.uniprot.org/sparql">http://sparql.uniprot.org/sparql</a></td>
</tr>
<tr>
<td>DisGeNet</td>
<td><a href="http://rdf.disgenet.org/sparql/">http://rdf.disgenet.org/sparql/</a></td>
</tr>
<tr>
<td>MeSH</td>
<td><a href="http://id.nlm.nih.gov/mesh/sparql">http://id.nlm.nih.gov/mesh/sparql</a></td>
</tr>
<tr>
<td>Orphanet</td>
<td><a href="https://www.orpha.net/sparql">https://www.orpha.net/sparql</a></td>
</tr>
</tbody>
</table>
The DO-KB’s SPARQL Sandbox provides programmatic access to the DO-KB Knowledgebase for exploring DO diseases, disease features, cross references and mechanisms, and exploring disease-data connections to other SPARQL endpoints.

**Choose a Query:**
1. All diseases with their parents (identified by label)
2. Count of diseases in each branch (diseases may be in multiple branches)
3. Syndromes (branch search; returns IRI, ID and label)
4. Diseases with MeSH cross-references
5. Diseases with OMIM cross-references
6. Diseases with PubMed sources (limited to first 10 results)
7. Diseases in DO_infectious_disease_slim with their parents and branches
8. Diseases with their definitions and count of logical axioms
9. Diseases with their Evidence and Conclusion Ontology (ECO) reference types (queries the doi2merged-owl graph)
10. Count of classes imported from other ontologies (queries the doi2merged-owl graph)
11. All diseases with equivalent class axioms (axioms are formatted for readability and may not be displayed in full; queries the doi2merged-owl graph)
12. Extract the disease and supporting import class hierarchies using subclassOf relationships (executes a CONSTRUCT query that returns RDF triples)

**Query execution time:** 2.22 second(s).

<table>
<thead>
<tr>
<th>id</th>
<th>label</th>
<th>OMIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOID:0050120</td>
<td>hemophagocytic lymphohistiocytosis</td>
<td>OMIM:267700</td>
</tr>
<tr>
<td>DOID:0050156</td>
<td>idiopathic pulmonary fibrosis</td>
<td>OMIM:178500</td>
</tr>
<tr>
<td>DOID:0050158</td>
<td>desquamative interstitial pneumonia</td>
<td>OMIM:263000</td>
</tr>
<tr>
<td>DOID:0050167</td>
<td>autoimmune polyendocrine syndrome type 1</td>
<td>OMIM:240300</td>
</tr>
</tbody>
</table>

**Coming Soon:**
- Federated queries to UniProt, MeSH, Protein Ontology

- Select a SPARQL query to retrieve ML-ready datasets
- SELECT (table formatted results) or CONSTRUCT (RDF results) queries
- Edit queries to refine your search
- Suggest new SPARQL queries to DO-KB
- Utilize the DO-KB SPARQL Endpoint to build your own federated queries
Structuring a SPARQL Query

```sparql
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX oboInOwl: <http://www.geneontology.org/formats/oboInOwl#>

SELECT ?id ?label ?parent
WHERE {
  ?class a owl:Class .
  ?class oboInOwl:id ?id .
  ?class rdfs:label ?label .
  ?class rdfs:subClassOf ?parent_class .
}
```