

Module 3 – Ontologies

Ontologies and Protégé

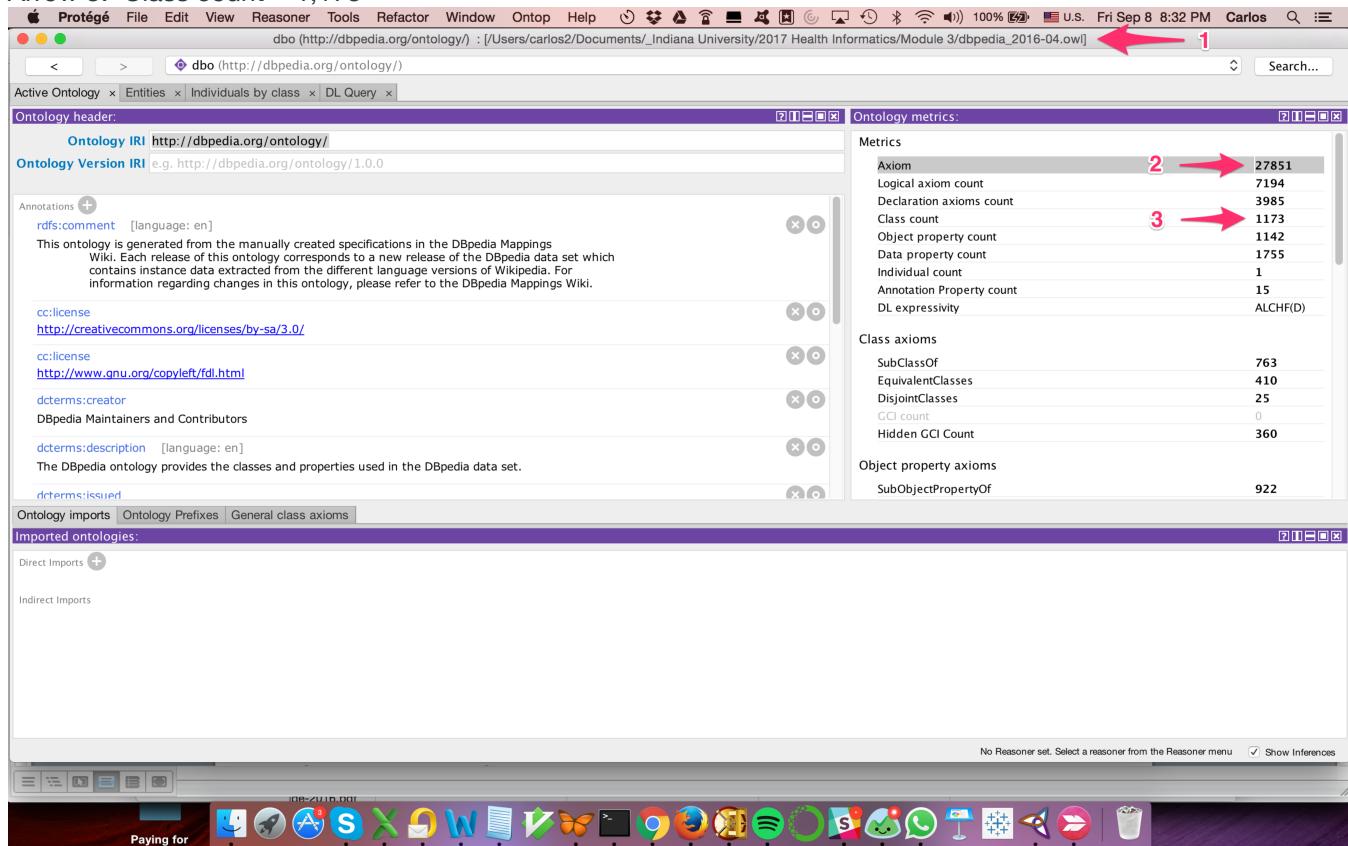
1. Download and install Protégé from <http://protege.stanford.edu/>. Protégé is a "free, open-source ontology editor and framework for building intelligent systems". Protégé can be used to browse and explore ontologies, which is the overall task in this assignment.
2. Download the latest DBpedia ontology in OWL format, from <http://wiki.dbpedia.org>. "DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web." Load the ontology into Protégé.
3. How many classes and axioms are in the ontology? In general an axiom is any assertion about a class, such as a SubClassOf relationship. Describe at least one class and one axiom in your report, with illustrative screen shots.

Screenshot 1 of 2:

Arrow 1: Shows DBpedia (in OWL format) has been loaded into Protégé

Arrow 2: Number of axioms = 27,851

Arrow 3: Class count = 1,173



DSDHT FALL 2017

Carlos Sathler (cssathler@gmail.com)

Screenshot 2 of 2:

Arrow 1: Shows selection of class MouseGene. In the hierarchy window pane we see MouseGene is a “gene”; a gene is a “Biomolecule”, and a Biomolecule is a “Thing”.

Arrow 2: Show the MouseGene description, which confirms what we see on the hierarchy pane (SubClass Of information). There is no equivalent class to Mousegene, no general class axiom, instance, target for key, disjoint class specification or disjoint union information. Additionally, the top right window panel shows annotations for class “MouseGene”, specifically the class label in English and several other languages.

Arrow 3: Show one axiom for class MouseGene: “MouseGene” is a subclass of class “gene”.

The screenshot shows the Protégé 5.1.0 interface. The left pane is the "Class hierarchy" for the class `MouseGene`. The right pane contains the "Annotations" and "SubClass Of" sections for `MouseGene`.

- Annotations:**
 - `rdfs:label` [language: en] MouseGene
 - `rdfs:label` [language: ga] géin luiche
 - `rdfs:label` [language: de]
- Description:** MouseGene (highlighted by a pink arrow labeled 2)
- SubClass Of:** gene (highlighted by a pink arrow labeled 3)

Arrows numbered 1, 2, and 3 point to the highlighted areas in the screenshot.

DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)

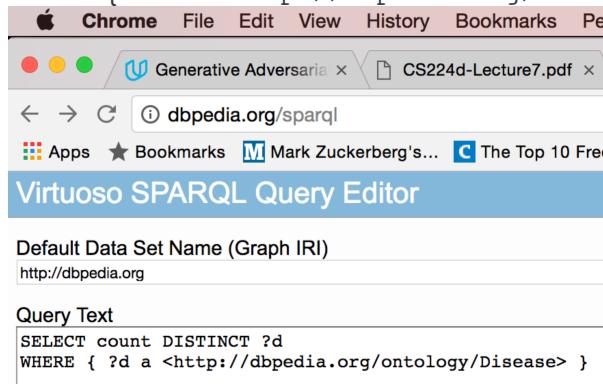
4. One DBpedia class relevant to health informatics is "disease" with URI <http://dbpedia.org/ontology/Disease>. Go to the DBpedia SPARQL endpoint <http://dbpedia.org/sparql> and run the SPARQL query below. How many diseases are found?

```
SELECT DISTINCT ?d
WHERE { ?d a <http://dbpedia.org/ontology/Disease> }
```

Screenshot 1 of 4:

First I ran a count of distinct values:

```
SELECT count DISTINCT ?d
WHERE { ?d a <http://dbpedia.org/ontology/Disease> }
```



Default Data Set Name (Graph IRI)
<http://dbpedia.org>

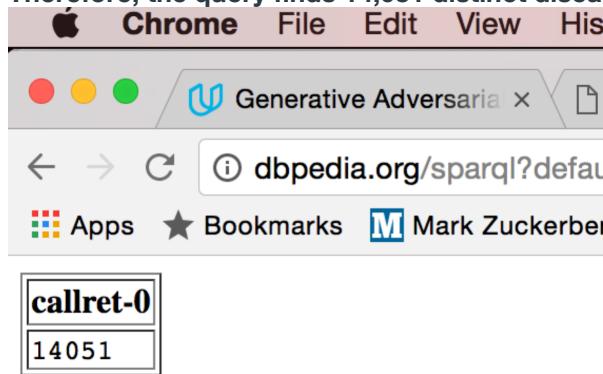
Query Text

```
SELECT count DISTINCT ?d
WHERE { ?d a <http://dbpedia.org/ontology/Disease> }
```

Screenshot 2 of 4:

Here's a screenshot of the result of the above query.

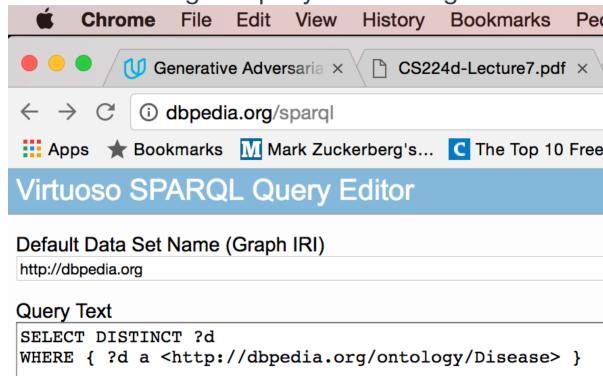
Therefore, the query finds 14,051 distinct diseases.



callret-0
14051

Screenshot 3 of 4:

Now for the original query in the assignment:



Default Data Set Name (Graph IRI)
<http://dbpedia.org>

Query Text

```
SELECT DISTINCT ?d
WHERE { ?d a <http://dbpedia.org/ontology/Disease> }
```

DSDHT FALL 2017

Carlos Sathler (cssathler@gmail.com)

Screenshot 4 of 4:

Results:

The screenshot shows a list of approximately 50 DBpedia resource URIs, each starting with <http://dbpedia.org/resource/>. The list includes terms like 'Acrochordon', 'Agoraphobia', 'Botulism', and various diseases and symptoms. The browser interface includes a toolbar at the top with icons for file operations, a search bar, and a list of open tabs.

- Each disease found should correspond with a Wikipedia "Infobox". Confirm this is true for a few diseases. Describe the role of the DBPedia ontology relative to DBPedia and Wikipedia.

A few infoboxes:

<p>Acrochordon</p>  <p>Several acrochorda in the skin of the lower neck, soft consistency, the bottom acrochordon taking a pedunculated shape</p> <p>Classification and external resources</p> <ul style="list-style-type: none"> Specialty Dermatology ICD-10 L91.8 (congenital Q82.8) ICD-9-CM 701.9 OMIM 109400 DiseasesDB 33273 MedlinePlus 000848 Patient UK Acrochordon <p>[edit on Wikidata]</p>	<p>Agoraphobia</p>  <p>An ancient agora in Delos, Greece. One of the public spaces after which the condition is named.</p> <p>Classification and external resources</p> <ul style="list-style-type: none"> Specialty Psychiatry Symptoms Anxiety in situations perceived to be unsafe, panic attacks^{[1][2]} Complications Depression, substance use disorder^[1] Duration > 6 months^[1] Causes Genetic and environmental factors^[1] Risk factors Family history, stressful event^[1] Similar conditions Separation anxiety, posttraumatic stress disorder, major depressive disorder^[1] Treatment Cognitive behavioral therapy^[3] Prognosis Resolution in half with treatment^[4] Frequency 1.7% of adults^[1] <p>[edit on Wikidata]</p>	<p>Botulism</p>  <p>A 14-year-old with botulism. Note the weakness of his eye muscles and the drooping eyelids in the adjacent image, and the large and non-moving pupils in the right image. This youth was fully conscious.</p> <p>Classification and external resources</p> <ul style="list-style-type: none"> Specialty Infectious disease, gastroenterology Symptoms Weakness, trouble seeing, feeling tired, trouble speaking^[1] Causes Clostridium botulinum^[1] Diagnostic method Finding the bacteria or its toxin^[1] Similar conditions Myasthenia gravis, Guillain–Barré syndrome, Amyotrophic lateral sclerosis, Lambert Eaton syndrome^[2] Prevention Proper food preparation, no honey for children less than one^[1] Treatment Antitoxin, antibiotics, mechanical ventilation^[1] Prognosis ~7.5% risk of death^[1] <p>[edit on Wikidata]</p>
--	--	--

DSDHT FALL 2017

Carlos Sathler (cssathler@gmail.com)

The role of the DBpedia ontology relative to DBpedia and Wikipedia:

Per the DBpedia About page (<http://wiki.dbpedia.org/about>) the “DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link the different data sets on the Web to Wikipedia data.” The DBpedia implements a knowledge base and it makes it available for access to the general public through a number of tools, such as DBpedia Lookup API’s, SPARQL, and the DBpedia Ontology. According to the DBpedia Ontology page (<http://wiki.dbpedia.org/services-resources/ontology>) the “DBpedia Ontology is a shallow, cross-domain ontology, which has been manually created based on the most commonly used infoboxes within Wikipedia.”

Therefore, the role of the DBpedia ontology relative to DBpedia and Wikipedia is to serve as a tool to allow access to the DBpedia knowledge base, which in its turn, constitutes a knowledge base for Wikipedia content.

Ontology Lookup Services

1. Go to the EBI Ontology Lookup Service at <http://www.ebi.ac.uk/ols/index> ([Links to an external site.](#)). Enter “asthma” in the search box. How many results are found? How many results are found when you filter on property? Explain the result. How many results are found when you filter on individual? In What ontology are the individuals found? How many results are found when you filter on ontology? Explain the result.

How many results are found?

376

The screenshot shows a web browser window with the URL www.ebi.ac.uk/ols/search?q=asthma. The page title is "Ontology Lookup Service". The search bar contains "asthma". Below the search bar are two checkboxes: "Exact match" (unchecked) and "Obsolete terms" (unchecked). A "Filters" button is visible. To the right, the text "Search results for *asthma*" is displayed, followed by "Showing 1 to 10 of 376 results". The first result is "Asthma [NCIT_C28397]", with a link to http://purl.obolibrary.org/obo/NCIT_C28397. It is described as "A chronic respiratory disease manifested as difficulty breathing due to the narrowing of bronchial passageways." and is associated with "Ontology: NCI Thesaurus OBO Edition [NCIT]". The second result is "Asthma [HP:0002099]", with a link to http://purl.obolibrary.org/obo/HP_0002099. It is described as "Asthma is characterized by increased responsiveness of the tracheobronchial tree to multiple stimuli, leading to narrowing of the air passages with resultant dyspnea, cough, and wheezing." and is associated with "Ontology: human phenotype ontology [HP]". There is also a note "Also in: OHMI [OHMI] DOID [DOID]". At the bottom of the page, there is a footer with various icons and the text "Paying for".

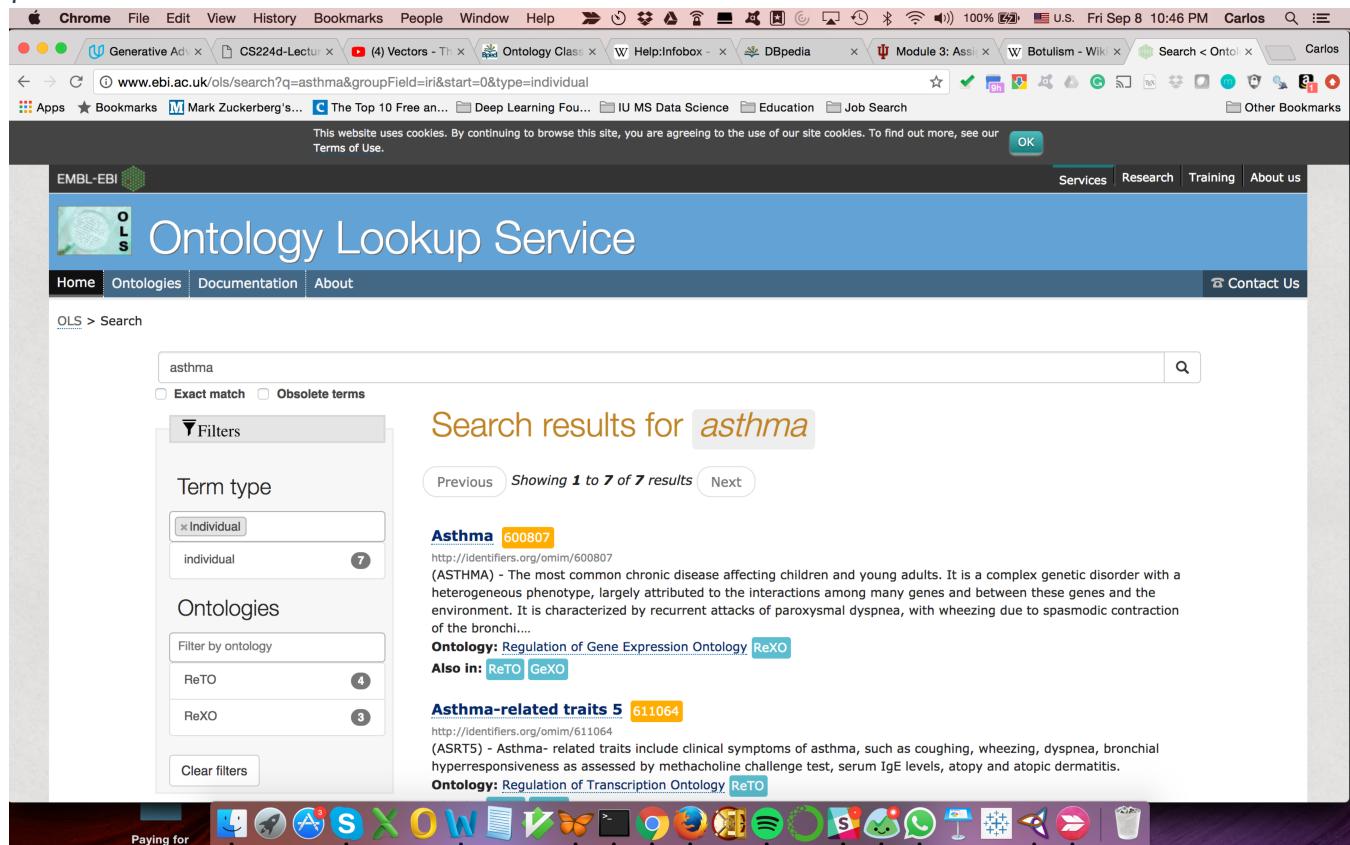
How many results are found when you filter on property? Explain the result.

Zero, because there are no ontologies that have “asthma” defined as a property.

DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)

How many results are found when you filter on individual?

7



The screenshot shows a web browser window with multiple tabs open. The active tab is 'www.ebi.ac.uk/ols/search?q=asthma&groupField=iri&start=0&type=individual'. The page title is 'Ontology Lookup Service'. The search bar contains 'asthma'. A sidebar on the left titled 'Filters' shows 'Term type' set to 'Individual' (with 7 results) and 'Ontologies' showing 'ReTO' (4 results) and 'ReXO' (3 results). The main content area displays search results for 'asthma':

- Asthma [600807]**
http://identifiers.org/omim/600807
(ASTHMA) - The most common chronic disease affecting children and young adults. It is a complex genetic disorder with a heterogeneous phenotype, largely attributed to the interactions among many genes and between these genes and the environment. It is characterized by recurrent attacks of paroxysmal dyspnea, with wheezing due to spasmodic contraction of the bronchi....
Ontology: Regulation of Gene Expression Ontology ReXO
Also in: ReTO GeXO
- Asthma-related traits 5 [611064]**
http://identifiers.org/omim/611064
(ASRT5) - Asthma- related traits include clinical symptoms of asthma, such as coughing, wheezing, dyspnea, bronchial hyperresponsiveness as assessed by methacholine challenge test, serum IgE levels, atopy and atopic dermatitis.
Ontology: Regulation of Transcription Ontology ReTO

In What ontology are the individuals found?

4 results are found in ontology ReTO and 3 results are found in ontology ReXO

How many results are found when you filter on ontology? Explain the result.

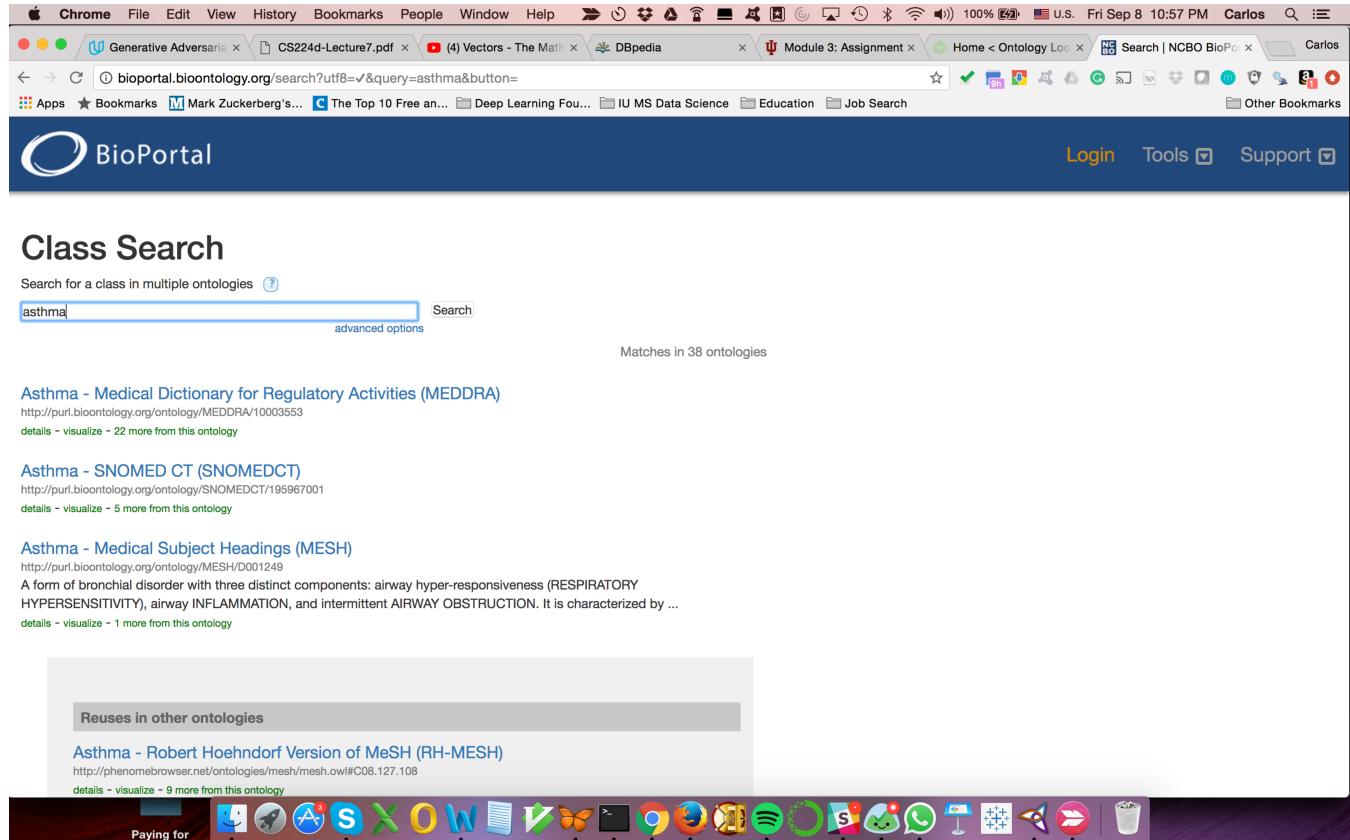
Zero, because there are no ontologies called "asthma".

DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)

2. Go to the National Center for BioMedical Ontology BioPortal site: <http://bioportal.bioontology.org/> ([Links to an external site.](#))[Links to an external site..](#) Enter “asthma” into the Search all ontologies search box. How many results are found? Click on the MeSH result, then expand the Asthma term. How many subheadings are under asthma? What are they? Asthma appears directly under what subject heading?

How many results are found?

Found 38



The screenshot shows a Google Chrome browser window with the BioPortal homepage loaded. The address bar shows the query: `bioportal.bioontology.org/search?utf8=%E2%8B%85&query=asthma&button=`. The main content area is titled "Class Search" and contains a search bar with "asthma" typed in. Below the search bar, it says "Matches in 38 ontologies". Three ontology results are listed:

- Asthma - Medical Dictionary for Regulatory Activities (MEDDRA)**
http://purl.bioontology.org/ontology/MEDDRA/10003553
[details](#) - [visualize](#) - 22 more from this ontology
- Asthma - SNOMED CT (SNOMEDCT)**
http://purl.bioontology.org/ontology/SNOMEDCT/195967001
[details](#) - [visualize](#) - 5 more from this ontology
- Asthma - Medical Subject Headings (MESH)**
http://purl.bioontology.org/ontology/MESH/D001249
A form of bronchial disorder with three distinct components: airway hyper-responsiveness (RESPIRATORY HYPERSENSITIVITY), airway INFLAMMATION, and intermittent AIRWAY OBSTRUCTION. It is characterized by ...
[details](#) - [visualize](#) - 1 more from this ontology

Below the ontology results, there is a section titled "Reuses in other ontologies" which lists "Asthma - Robert Hoehdorf Version of MeSH (RH-MESH)" with its URL: <http://phenomebrowser.net/ontologies/mesh/mesh.owl#C08.127.108>. At the bottom of the page, there is a "Paying for" banner featuring various payment method icons.

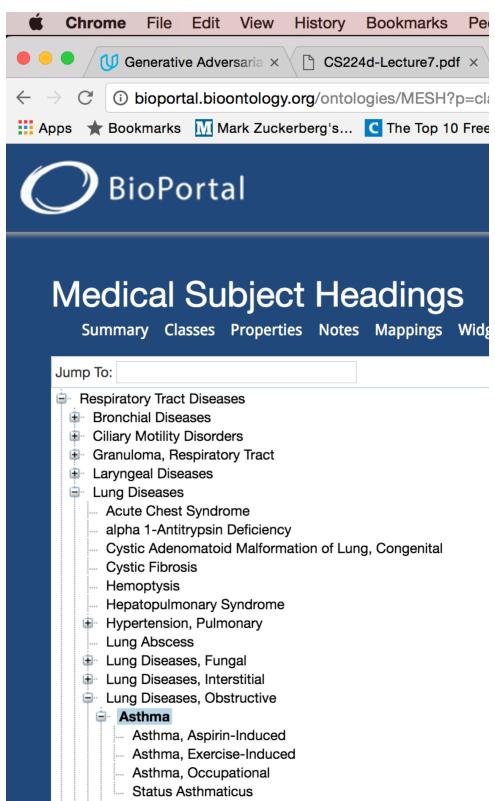
Click on the MeSH result, then expand the Asthma term. How many subheadings are under asthma? What are they? Asthma appears directly under what subject heading?

Please see screenshot below.

There are 4 subheadings under asthma. They are: (1) Asthma, Aspirin-Induced; (2) Asthma, Exercise-Induced; (3) Asthma, Occupational and (4) Status Asmathicus.

Asthma appears directly under subheading “Lung Diseases, Obstructive”

DSDHT FALL 2017
Carlos Sathler (cssathler@gmail.com)



The screenshot shows a Google Chrome browser window. The address bar displays "Generative Adversarial" and "CS224d-Lecture7.pdf". Below the address bar, the menu bar includes File, Edit, View, History, Bookmarks, and Print. The main content area is the BioPortal interface for Medical Subject Headings. The header features the BioPortal logo and the title "Medical Subject Headings". Below the title are navigation links: Summary, Classes, Properties, Notes, Mappings, Widgets, and Help. A "Jump To:" search bar is present. On the left, a tree view of medical subject headings is shown, with "Asthma" expanded to show sub-categories like "Asthma, Aspirin-Induced", "Asthma, Exercise-Induced", "Asthma, Occupational", and "Status Asthmaticus".

- Respiratory Tract Diseases
 - Bronchial Diseases
 - Ciliary Motility Disorders
 - Granuloma, Respiratory Tract
 - Laryngeal Diseases
 - Lung Diseases
 - Acute Chest Syndrome
 - alpha 1-Antitrypsin Deficiency
 - Cystic Adenomatoid Malformation of Lung, Congenital
 - Cystic Fibrosis
 - Hemoptysis
 - Hepatopulmonary Syndrome
 - Hypertension, Pulmonary
 - Lung Abscess
 - Lung Diseases, Fungal
 - Lung Diseases, Interstitial
 - Lung Diseases, Obstructive
 - Asthma**
 - Asthma, Aspirin-Induced
 - Asthma, Exercise-Induced
 - Asthma, Occupational
 - Status Asthmaticus