



Tarefa Trabalho 04

XQuery

Dupla LPT:

Gabriel Henrique Rosa Oswaldo - 172185

Lucas Silva Lopes Do Carmo - 202110

Dupla JSS:

Gabriel De Alcantara Bomfim Silveira - 197244

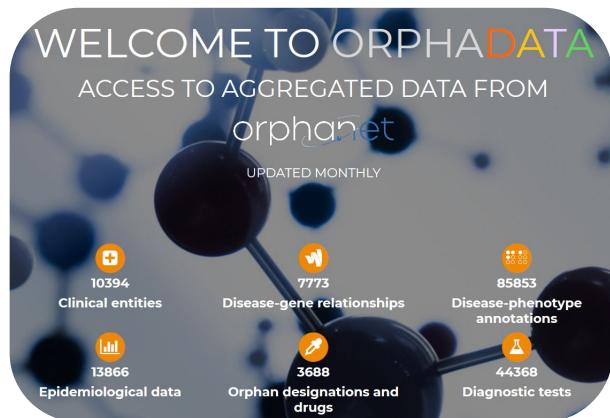
Vitor Coppo Ferreira - 206956

Orphadata

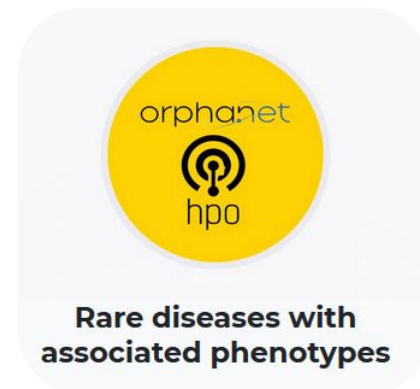
Rare Disease-Phenotype DB



- <http://www.orphadata.org/cgi-bin/index.php>



- http://www.orphadata.org/data/xml/en_product4_HPO.xml





```
1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <JDBOR date="2019-10-01 04:09:38" version="1.2.11 / 4.1.6 [2018-04-12] (orientdb version)"
  copyright="Orphanet (c) 2019">
3 <Availability>
4 <Licence>
5   <FullName lang="en">Creative Commons Attribution 4.0 International</FullName>
6   <ShortIdentifier>CC-BY-4.0</ShortIdentifier>
7   <LegalCode>https://creativecommons.org/licenses/by/4.0/legalcode</LegalCode>
8 </Licence>
9 </Availability>
10 <DisorderList count="3702">
11 <Disorder id="17601">
12   <OrphaNumber>58</OrphaNumber>
13   <Name lang="en">Alexander disease</Name>
14   <HPODisorderAssociationList count="61">
15     <HPODisorderAssociation id="2">
16       <HPO id="2">
17         <HPOId>HP:0000218</HPOId>
18         <HPOTerm>High palate</HPOTerm>
19       </HPO>
20       <HPOFrequency id="28426">
21         <OrphaNumber>453313</OrphaNumber>
22         <Name lang="en">Occasional (29-5%)</Name>
23       </HPOFrequency>
24       <DiagnosticCriteria/>
25     </HPODisorderAssociation>
26     <HPODisorderAssociation id="3">
27       <HPO id="3">
28         <HPOId>HP:0000238</HPOId>
29         <HPOTerm>Hydrocephalus</HPOTerm>
30       </HPO>
31       <HPOFrequency id="28426">
32         <OrphaNumber>453313</OrphaNumber>
33         <Name lang="en">Occasional (29-5%)</Name>
34       </HPOFrequency>
35       <DiagnosticCriteria/>
36     </HPODisorderAssociation>
37   </HPODisorderAssociationList>
38 </Disorder>
39 </DisorderList>
```

Doença rara

Nome da doença rara

Distúrbio associado
a doença

Nome do distúrbio

Frequência com que o
distúrbio se manifesta
na doença



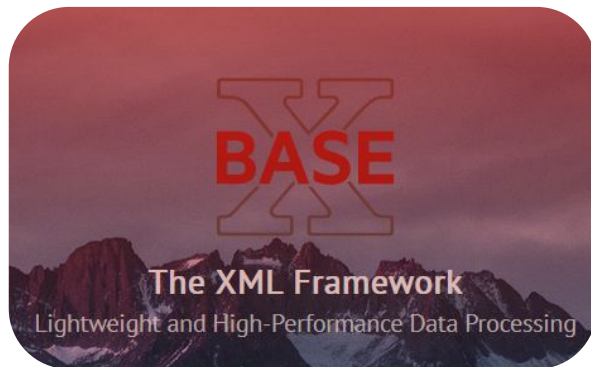
Por quê o modelo hierárquico?

- Quantidade variável de nós filhos (sintomas)
- Permite queries avançadas
- Grande quantidade de camadas
- Escalabilidade



Preparação dos Dados

- try.zorba.io/ não suportou a quantidade de dados
- Usamos então o [BaseX](#)
 - XML framework
 - XQuery 3.1 processor





XQueries

Consultas a fonte de dados



XQuery 01

Começamos com algumas consultas simples.

- Quantidade de doenças raras cadastradas na base
 - Resultado: **3702**

```
1 for $c in //DisorderList
2 return count($c/Disorder)
```



XQuery 02

- Quantidade de distúrbios associados a doença rara com o id especificado
 - Resultado: **61**

```
1 for $c in //Disorder[@id=2]//HPODisorderAssociationList
2 return count($c/HPODisorderAssociation)
```


XQuery 03



- Lista formatada de distúrbios com a frequência acima de “Frequente”, associadas a doença rara com o id especificado

```
1 for $c in //Disorder[@id=2]//HPODisorderAssociation
2 where $c/HPOFrequency[Name="Frequent (79-30%)"] or
3       $c/HPOFrequency[Name="Very frequent (99-80%)"] or
4       $c/HPOFrequency[Name="Obligate (100%)"]
5 return <sintoma>
6         <nome>{data($c//HPOTerm)}</nome>
7         <frequencia>{data($c/HPOFrequency/Name)}</frequencia>
8     </sintoma>
```

```
<sintoma>
  <nome>Macrocephaly</nome>
  <frequencia>Very frequent (99-80%)</frequencia>
</sintoma>
<sintoma>
  <nome>Abnormality of eye movement</nome>
  <frequencia>Frequent (79-30%)</frequencia>
</sintoma>
<sintoma>
  <nome>Ptosis</nome>
  <frequencia>Frequent (79-30%)</frequencia>
</sintoma>
```



XQuery 04

```
1 for $c in //Disorder
2 where $c//HPODisorderAssociation/HPO[HPOTerm="Macrocephaly"] and
3     ($c//HPODisorderAssociation//HPOFrequency[Name="Very frequent (99-80%)"] or
4     $c//HPODisorderAssociation//HPOFrequency[Name="Frequent (79-30%)"] or
5     $c//HPODisorderAssociation//HPOFrequency[Name="Obligate (100%)"])
6 return data($c/Name)
```

- Lista de doenças que possuem o sintoma especificado com frequência acima de "Frequent (79-30%)"

Multiple epiphyseal dysplasia, Al-Gazali type
Alexander disease
Alpha-mannosidosis
Multiple sulfatase deficiency
Canavan disease
Sandhoff disease
X-linked intellectual disability, Nascimento type
Fragile X syndrome
Mucopolysaccharidosis type 2
Mucopolysaccharidosis type 1



XQuery 05

- Lista doenças que possam dois sintomas ao mesmo tempo com uma frequência acima de ocasional

```
1 for $c in (//Disorder),
2   $d in (//Disorder)
3 where $c//HPODisorderAssociation/HPO[HPOTerm="Macrocephaly"] and
4       $d//HPODisorderAssociation/HPO[HPOTerm="Nystagmus"] and
5       ($c[@id] = $d[@id])
6 and
7       (($c//HPODisorderAssociation/HPOFrequency[Name="Frequent (79-30%)"] or
8         $c//HPOFrequency[Name="Very frequent (99-80%)"] or
9         $c//HPOFrequency[Name="Obligate (100%)"]) and
10        ($d//HPOFrequency[Name="Frequent (79-30%)"] or
11         $d//HPOFrequency[Name="Very frequent (99-80%)"] or
12         $d//HPOFrequency[Name="Obligate (100%)"]))
13
14 return data($c/Name)
```

Alexander disease
Zellweger syndrome
Neonatal adrenoleukodystrophy
6q16 microdeletion syndrome
Borjeson-Forssman-Lehmann syndrome
Spinocerebellar ataxia-dysmorphism syndrome
Joubert syndrome with hepatic defect
Distal limb deficiencies-micrognathia syndrome
Cardiofaciocutaneous syndrome
Dysosteosclerosis



Obrigado!

Alguma pergunta?