
QUESTIONS FROM THE COURSES

Day 01: questions from the course.

Q1.1 Practice XML replace missing parts

```
<archi_book>

  <short_title>Architecture Now</short_title>

  <main_author>Jodidio, Philip</main_author>

  <ID isbn10="3822840912"/>

</archi_book>
```

Q1.2 Provide 10 first lines

Get 10 first lines of the five results for:

<http://www.wikidata.org/entity/Q23014205>
<http://www.wikidata.org/entity/Q23014205.json>
<http://www.wikidata.org/entity/Q23014205.rdf>
<http://www.wikidata.org/entity/Q23014205.ttl>
<http://www.wikidata.org/entity/Q23014205.nt>

<!DOCTYPE html>
{"entities":{"Q23014205":{"pageid":25028548,"ns":0,"title":"Q23014205","lastrevid":872123137,"modified":"2019-03-02T12:03:45Z","type":"item","id":"Q23014205","labels":{"fr":
<?xml version="1.0"?>
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<<https://www.wikidata.org/wiki/Special:EntityData/Q23014205>> <<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>> <<http://schema.org/Dataset>> .

Q1.3 DBpedia

1. Find "London" on DBpedia.org; e.g. Google: "london site:dbpedia.org"
make sure you are on the English chapter (dbpedia.org) as there are many others (fr.dbpedia.org, de.dbpedia.org)
2. Find dbp:populationDonym and give its value
3. Find rdf:type and click on value yago:WikicatCapitalsInEurope
4. Find "Vienna" and get its URI
(careful: with content negotiation and redirection, the URL of the page you are currently viewing may be different from the URI of the resource it describes)
5. Access to Vienna and find its native name?

Londoner

<http://dbpedia.org/class/yago/WikicatCapitalsInEurope>

<http://dbpedia.org/resource/Vienna>

<http://dbpedia.org/page/Vienna>

Q1.4 WHO.IS?

1. contact for inria.fr
2. contact for fabien.info
3. contact for lemonde.fr

Institut National de Recherche en Informatique et en Automatique

Not available

SOCIETE EDITRICE DU MONDE

Q1.5 CURL

4. Ten first lines:

```
curl -o Paris.html -L -H "Accept: text/html" http://dbpedia.org/resource/Paris
```

```
curl -o Paris-rdf-xml.txt -L -H "Accept: application/rdf+xml" http://dbpedia.org/resource/Paris
```

2. Ten first lines for HTML and RDF <http://ns.inria.fr/fabien.gandon#me>
3. Ten first lines for HTML and RDF for 'Vienna' on Dbpedia
4. Ten first lines for the "URI of the name of Victor Hugo" in the Library of Congress:
<http://id.loc.gov/authorities/names/n79091479>
5. Ten first lines for HTML and RDF
<http://purl.uniprot.org/uniprot/P43121>
6. What is the topic and format of data obtained with
`curl -o json.txt -L -H "Accept: application/json" https://www.wikidata.org/wiki/Special:EntityData/Q551861`
7. What is the topic and format of data obtained with
`curl -o turtle.txt -L -H "Accept: text/turtle" http://dx.doi.org/10.1007/3-540-45741-0_18`

<pre><?xml version="1.0" encoding="UTF-8" ?> <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN" "http://www.w3.org/MarkUp/DTD/xhtml- rdfa-1.dtd"> <html xmlns="http://www.w3.org/1999/xhtml" xmlns:dbpprop="http://dbpedia.org/property/" xmlns:foaf="http://xmlns.com/foaf/0.1/" version="XHTML+RDFa 1.0" xml:lang="en" ></pre>	<pre><?xml version="1.0" encoding="utf-8" ?> <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf- syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf- schema#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:dbo="http://dbpedia.org/ontology/" xmlns:dct="http://purl.org/dc/terms/" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:geo="http://www.w3.org/2003/01/geo/wg s84_pos#" xmlns:prov="http://www.w3.org/ns/prov#" ></pre>
<pre><?xml version="1.0" encoding="utf-8" ?> <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1- strict.dtd"> <html xmlns="http://www.w3.org/1999/xhtml"></pre>	<pre><?xml version='1.0' encoding='utf-8' ?> <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf- syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf- schema#" ></pre>

<pre> <head> <meta http-equiv="Content-Type" content="text/html;charset=utf-8" /> <title>FOAF profile of Fabien GANDON</title> </head> <body> <h1>FOAF profile of Fabien GANDON</h1> <p>You may have been redirected here by your browser.</p> </pre>	<pre> xmlns:foaf="http://xmlns.com/foaf/0.1/" xml:base="http://ns.inria.fr/fabien.gandon"> <foaf:PersonalProfileDocument rdf:about=""> <foaf:maker rdf:resource="#me"/> <foaf:primaryTopic rdf:resource="#me"/> </pre>
<pre> <?xml version="1.0" encoding="UTF-8" ?> <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN" "http://www.w3.org/MarkUp/DTD/xhtml- rdfa-1.dtd"> <html xmlns="http://www.w3.org/1999/xhtml" xmlns:dbpprop="http://dbpedia.org/property/" xmlns:foaf="http://xmlns.com/foaf/0.1/" version="XHTML+RDFa 1.0" xml:lang="en" > </pre>	<pre> <!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN"> <html><head> <title>406 Not Acceptable</title> </head><body> <h1>406 Not Acceptable</h1> <p>An appropriate representation of the requested resource Vienna could not be found on this server.</p> Available variant(s): Vienna , type text/html, charset UTF-8 </pre>
<pre> <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN" "http://www.w3.org/MarkUp/DTD/xhtml-rdfa-1.dtd"> <html version="XHTML+RDFa 1.0" xmlns="http://www.w3.org/1999/xhtml" xmlns:madsrdf="http://www.loc.gov/mads/rdf/v1#" xmlns:ri="http://id.loc.gov/ontologies/RecordInfo#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf- schema#" xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:skosxl="http://www.w3.org/2008/05/skos-xl#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:cs="http://www.w3.org/2003/06/sw-vocab-status/ns#" xmlns:dcterms="http://purl.org/dc/terms/"> </pre>	

<head>

<title>Hugo, Victor, 1802-1885 - LC Linked Data Service: Authorities and Vocabularies | Library of Congress</title>

<meta name="description" content=" The Linked Data Service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. Datasets available include LCSH, BIBFRAME, LC Name Authorities, LC Classification, MARC codes, PREMIS vocabularies, ISO language codes, and more."/>

<!DOCTYPE html SYSTEM "about:legacy-compat">

<html xmlns="http://www.w3.org/1999/xhtml" lang="en" xml:lang="en"><head><title>MCAM - Cell surface glycoprotein MUC18 precursor - Homo sapiens (Human) - MCAM gene & protein</title><meta content="IE=edge" http-equiv="X-UA-Compatible"/><meta content="text/html; charset=UTF-8" http-equiv="Content-Type"/><meta content="width=device-width, initial-scale=1" name="viewport"/><link href="/" rel="home"/><link href="https://creativecommons.org/licenses/by/4.0/" rel="license"/><link type="image/vnd.microsoft.icon" href="/favicon.ico" rel="shortcut icon"/><link href="/uniprot.min.css2020_01" type="text/css" rel="stylesheet"/><script type="text/javascript">

var BASE = '/';

var ua = window.navigator.userAgent;

var directory = (~ua.indexOf('MSIE ')) || ~ua.indexOf('Trident/')) === 0 ? "non-ie" : "ie";

</script><script src="/scripts/frontier/d3/d3.v3.min.js" type="text/javascript"></script><script src="/js-compr.js2020_01" type="text/javascript"></script><script type="text/javascript">

uniprot.namespace = 'uniprot';

uniprot.releasedate = '2020_01';

<?xml version='1.0' encoding='UTF-8'?>

<rdf:RDF xml:base="http://purl.uniprot.org/uniprot/" xmlns="http://purl.uniprot.org/core/" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:bibo="http://purl.org/ontology/bibo/" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:void="http://rdfs.org/ns/void#" xmlns:sd="http://www.w3.org/ns/sparql-service-description#" xmlns:faldo="http://biohackathon.org/resource/faldo#">

<owl:Ontology

rdf:about="http://purl.uniprot.org/uniprot/">

<owl:imports

rdf:resource="http://purl.uniprot.org/core/">

</owl:Ontology>

<rdf:Description

rdf:about="http://purl.uniprot.org/uniprot/P43121">

<rdf:type

rdf:resource="http://purl.uniprot.org/core/Protein"/>

<reviewed

rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean">true</reviewed>

<pre> </script><script type="text/javascript"> ; </pre>	<pre> <created rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1995-11-01</created> <modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2020-02-26</modified> </pre>
json.txt	json
turtle.txt	turtle

Q1.6 Recall five best practices of linked open data



On the web

Machine-readable data

Non-proprietary format

RDF standards

Linked RDF

Q1.7 Spotlight demo (<https://www.dbpedia-spotlight.org/demo/>)

Reproduce the demo:

1. Copy a text from Wikipedia (e.g. Muse Band page)
2. Find the DBpedia Spotlight service page
3. Paste the text and run the detection
4. Try with other texts and copy-paste one of the results you get.

Muse are an English rock band from Teignmouth, Devon, formed in 1994. The band consists of Matt Bellamy (lead vocals, guitar, keyboards), Chris Wolstenholme (bass guitar, backing vocals), and Dominic Howard (drums).