Convert your data to Linked Data with OpenRefine

These slides: http://bit.ly/dh-slides

Go to OpenRefine

http://refine.amp.ops.labs.vu.nl/

If you need your own instance:

- Get LODRefine; or
- Install OpenRefine and then the RDF extension

Download CSV example

https://tinyurl.com/ld4dhcsv

```
Rank;Country;Int
1;Qatar;131,063
2;Luxembourg;104,906
3;Macau;96,832
4;Singapore;90,249
5;Brunei Darussalam;83,513
6;Kuwait;72,675
7;Ireland;72,524
8;Norway;70,645
```

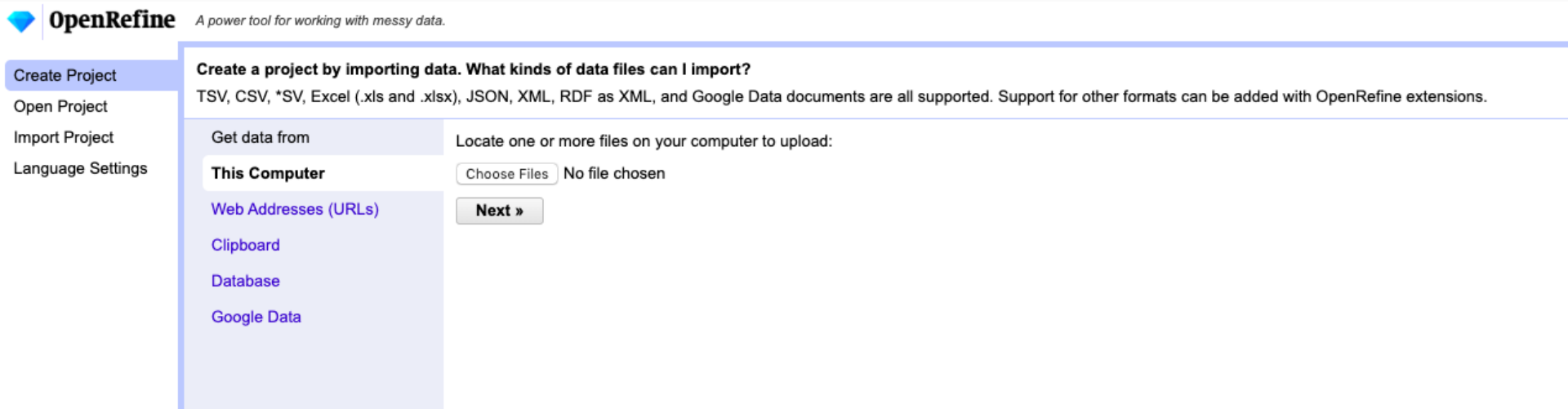
Today’s goal: turn that CSV into meaningful RDF (Linked Data)

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix example: <http://purl.org/collections/example/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

example:knud
  a foaf:Person ;
  foaf:name “Knud Möller”^^xsd:string ;
  foaf:knows example:victor ;
  foaf:topic_interest example:linked_data .

example:victor
  a foaf:Person ;
  foaf:name “Victor de Boer”^^xsd:string ;
  foaf:knows example:knud ;
  foaf:topic_interest example:linked_data ;
  foaf:knows example:truffel .
```

Create project



RDF Schema alignment

The RDF schema alignment skeleton below specifies how the RDF data that will get generated from your grid-shaped data. The cells in each record of your data will get placed into nodes within the skeleton. Configure the skeleton by specifying which column to substitute into which node.

Base URI: http://refine.amp.ops.labs.vu.nl/ Edit

RDF skeleton | RDF Preview

Available prefixes: rdf owl rdfs foaf Add Manage

(Row index) URI
Add type

property?
property?
property?
Add property

Rank Cell
Country Cell
Int Cell

Add another root node

Save

OK Cancel

“Subjects” or things we want to describe

What are the **main concepts**, entities, things of the real world… in that CSV that we want to describe as Linked Data?

(hint: look at the column names)

Predicates and reuse

https://lov.linkeddata.es is a **general catalog of RDF vocabularies** for reuse

For columns “Int”, “Rank” you can use:

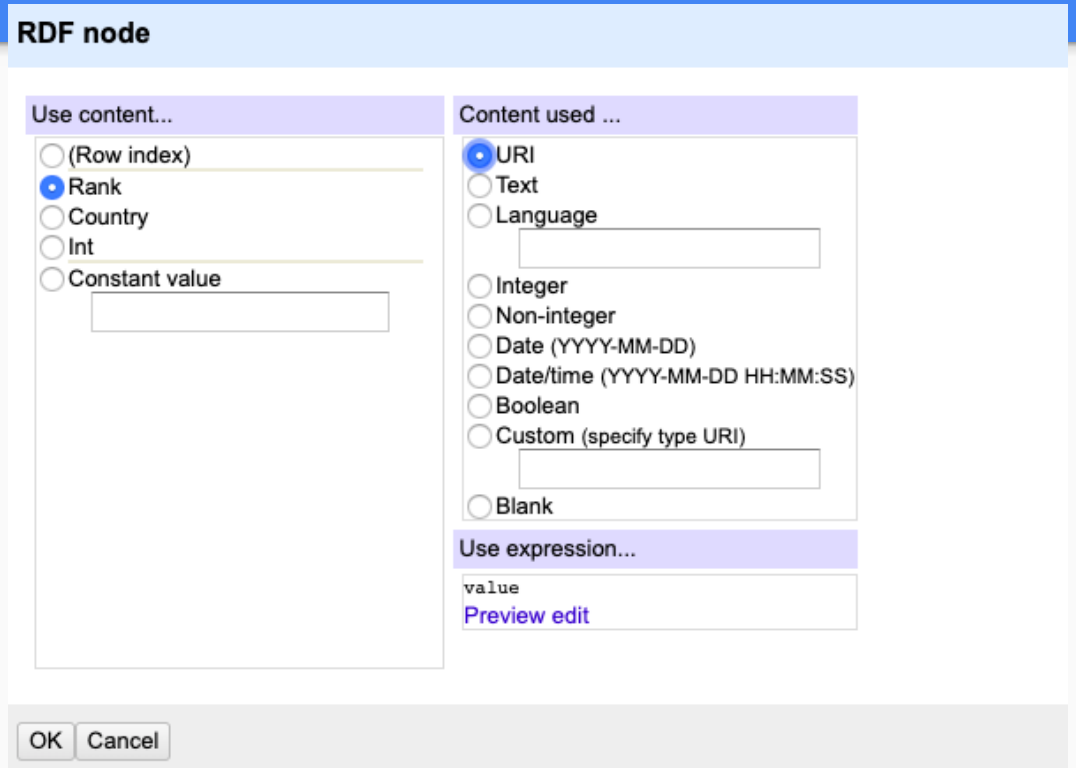
http://aims.fao.org/aos/geopolitical.owl#GDP

http://purl.org/linked-data/cube#order

For saying that you created those countries, you can use:

http://purl.org/dc/terms/creator

“Objects” as URIs or Literals



Other “root nodes” or subjects

- Used to describe independent entities
  - In other columns of your CSV
  - Elsewhere

Think about **other things** related to the “main concept” of the CSV you want to describe as Linked Data

(hint: who’s the author of the CSV statements? Where did you get it from? etc.)

More: subject types

rdf:type expresses class membership (or “type”)

More: external links

In Linked Data, besides describing everything as URIs and subject-predicate-object triples, **you need to link to the URIs of other people**

E.g. http://dbpedia.org/resource/Qatar describes exactly the same real world entity as your own

We do this using the **owl:sameAs** predicate

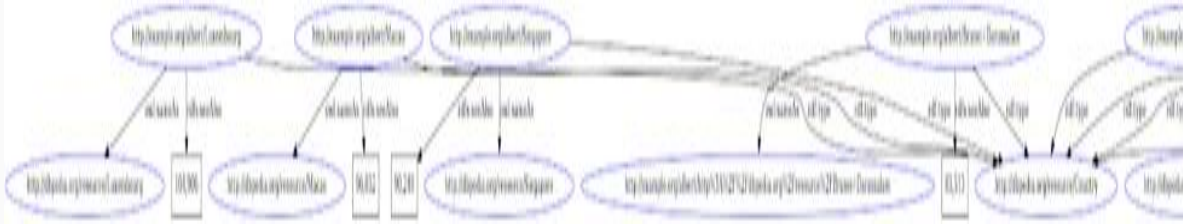
More: export, check output, visualize

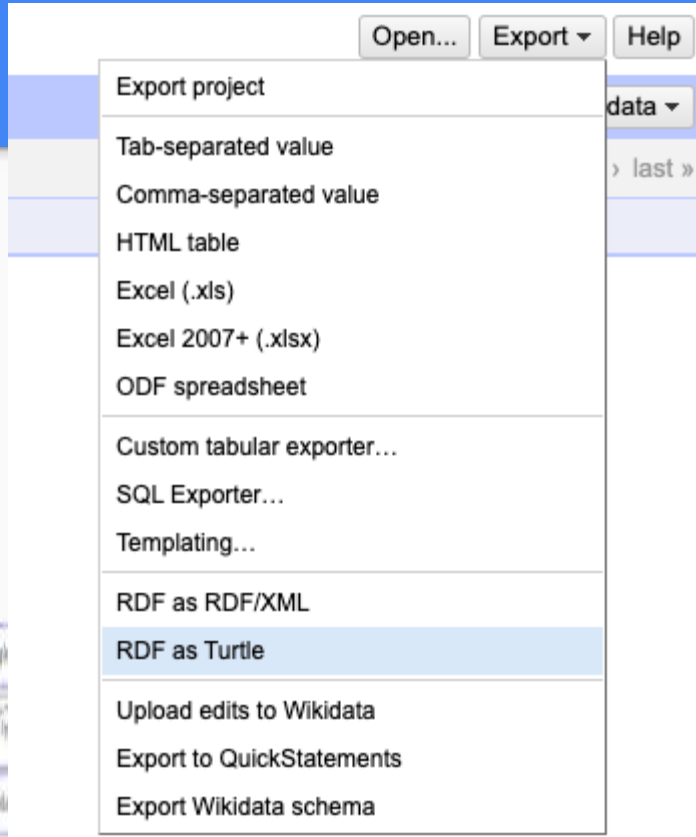
Export your RDF as Turtle

Open it with a text editor and examine it

Use http://www.easyrdf.org/converter to visualize it

(set “Turtle” as input, “SVG” as output, check “raw”)





Other possibilities

- Reconcile with DBpedia/Wikidata
- Add arbitrary SPARQL endpoints/files
- Namespace management
- Advanced GREL expressions
- Blank nodes (for “complex” yet unnamed concepts)

Thank you!

Go get your CSV/XLS/… data and quickly transform it into RDF with OpenRefine

Questions?