

Introduction to data migration for SNAP

Sebastian Rahtz

24 March 2014



Making it go SNAP

- 1 do you have **structured** information about ancient world persons?
- 2 is your data **machine-readable** (digitized)?
- 3 can you put it **online**?
- 4 are there **stable identifiers**?

Congratulations, you are SNAPable



Your data translation strategy: examples

SQL database you'll need a database report

XML you can make an XSLT transformation

CSV files you can make a Python/Perl/Java script

Word files Word .docx files are actually XML

RDF You can write SPARQL to construct new RDF



Tools to assist you

There are a *lot* of tools to help you write RDF. For example

- <http://www.w3.org/RDF/> is a dry but correct starting point
- <http://www.w3.org/wiki/ConverterToRdf> is rather technical list
- http://logd.tw.rpi.edu/technology/rdf_extension_google_refine talks about using Google Refine



Template 1: the Person

```
<http://www.lgpn.ox.ac.uk/id/V3a-40899#this> rdf:type  
lawd:Person .
```



Template 2: the person's simple properties

```
<http://www.lgpn.ox.ac.uk/id/V5a-35652>  
foaf:name "Ἀπολλώνιος"@grc ,  
"Apollwnios"@grc-Latn ,  
"Apollonios"@en ,  
ar@"ابولونيوس" .
```



Template 3: simple date

```
<http://www.lgpn.ox.ac.uk/id/V5a-35652> dc:date "0101/0200"
```

.



Template 4: place link

```
<http://www.trismegistos.org/person/1234#this>  
lawd:where <http://pleiades.stoa.org/places/246474> .
```



Template 5: place link as text

```
<http://www.example.com/person/789#this>  
lawd:where <http://www.example.com/place/4321> .  
<http://www.example.com/place/4321>  
rdf:type lawd:Place, cnt:ContentAsText ;  
cnt:chars "Chersakes"@en ,  
"Χέρσακες"@grc .
```



Template 6: hasName

```
<http://www.lgpn.ox.ac.uk/id/V5a-35652#this> lawd:hasName  
<http://www.lgpn.ox.ac.uk/nym/nAglwnilkh> .
```



Template 7: PersonalName

```
<http://www.lgpn.ox.ac.uk/nym/nAglwnikh>  
a          lawd:PersonalName ;  
lawd:primaryForm  "Άγλωνίκη"@grc , "Aglwnikh"@grc-Latn ;  
dc:publisher    <http://www.lgpn.ox.ac.uk/> .
```



Template 7: hasAttestation

```
<http://www.lgpn.ox.ac.uk/id/V5a-57002>  
lawd:hasAttestation  
<http://www.lgpn.ox.ac.uk/id/V5a-57002/personref/1>
```



Template 8: Attestation

```
<http://www.lgpn.ox.ac.uk/id/V1-1006/personref/1>  
a lawd:PersonAttestation ;  
dc:publisher <http://www.lgpn.ox.ac.uk/> ;  
cito:citesAsEvidence  
<http://www.lgpn.ox.ac.uk/id/V1-1006/ref/1> .
```



Template 9: Citation

```
<http://www.lgpn.ox.ac.uk/id/V1-1006/ref/1>  
a          lawd:Citation , cnt:ContentAsText ;  
cnt:chars "IG XII Suppl. p. 132 no. 796, 2" .
```



Template 10: Relationships

```
<http://www.trismegistos.org/person/1234#this>  
snap:hasMother  
<http://www.trismegistos.org/person/1235#this> ;  
snap:hasFather  
<http://www.trismegistos.org/person/1236#this> .
```



Template 11: alt identifiers

```
<http://www.trismegistos.org/person/1234#this>  
dc:identifer <http://viaf.org/viaf/18013086/> ,  
<http://dbpedia.org/page/Augustus> .
```



How will you know when you've succeeded?

- You check the its syntactically ok by using an RDF editor or by using an XML well-formedness checker if you prefer RDF XML
- You can check its legal RDF by using an online validator (eg <http://www.w3.org/RDF/Validator/> or local command-line tools like rdfcat (comes with Apache Jena)
- These tools can usually convert between RDF formats, which is a useful check
- You can check your RDF is consonant with SNAP by ... staring at it? something for us to work on
- You can't check its **true**

