

Linking Maritime Datasets to Dutch Ships and Sailors Cloud - Case studies on Archangelvaart and Elbing

J.A. Entjes

July 10th, 2015



Table of Contents

- Introduction
- Research Questions
- Approach and Methodology
- Evaluation
- Recommendations
- Conclusion and Discussion

Introduction

- Digital Humanities
- Linked Data
- Dutch Ships and Sailors

Digital Humanities

- Historians and computers do not match
- Possibilities Linked Data has to offer

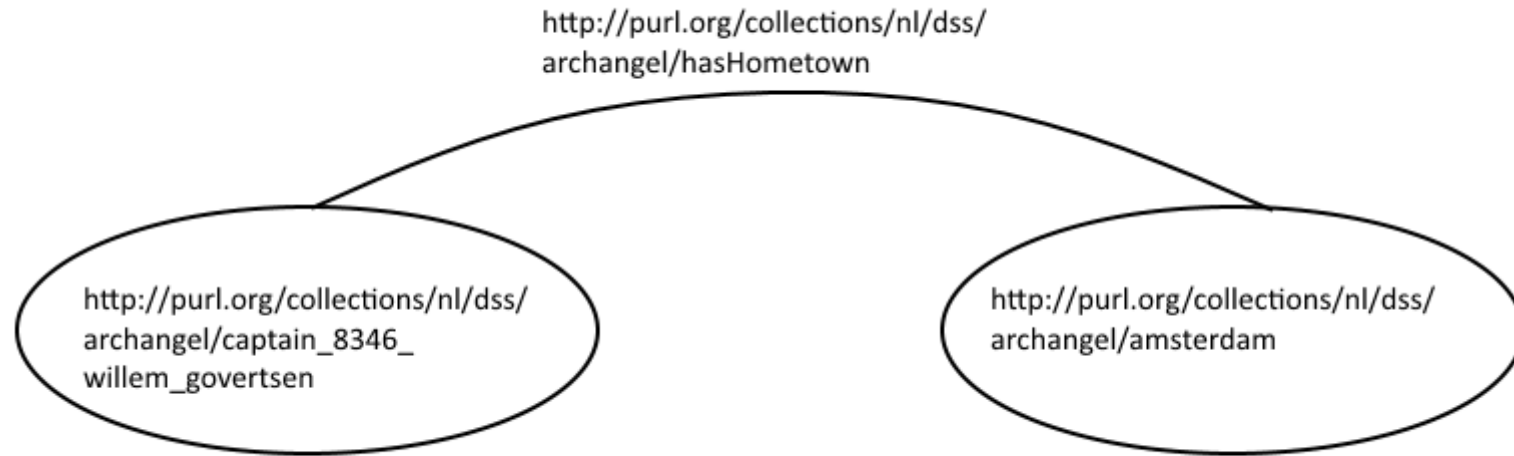
Schreibman et al., 2008
Meroño-Peñuela et al., 2013

Linked Data

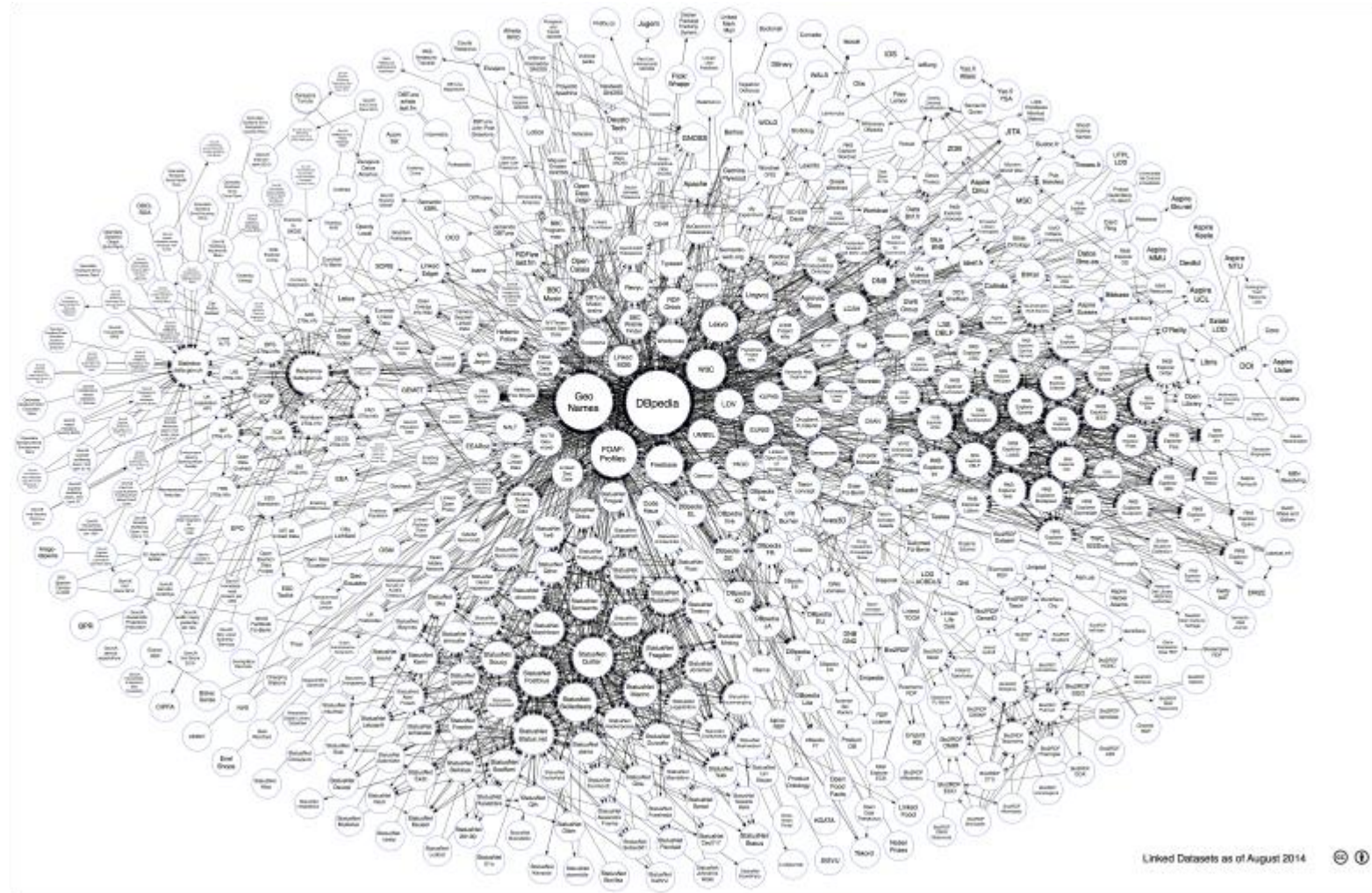
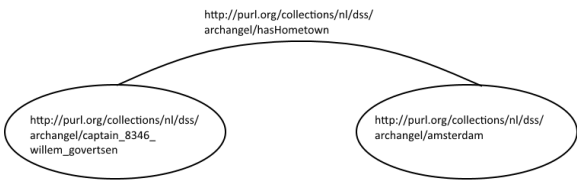
- RDF and SPARQL
- Uniform Resource Identifier
- Requirements for Linked Data
- Example

<http://www.w3.org/standards/semanticweb/>
Heath & Bizer, 2011

Linked Data



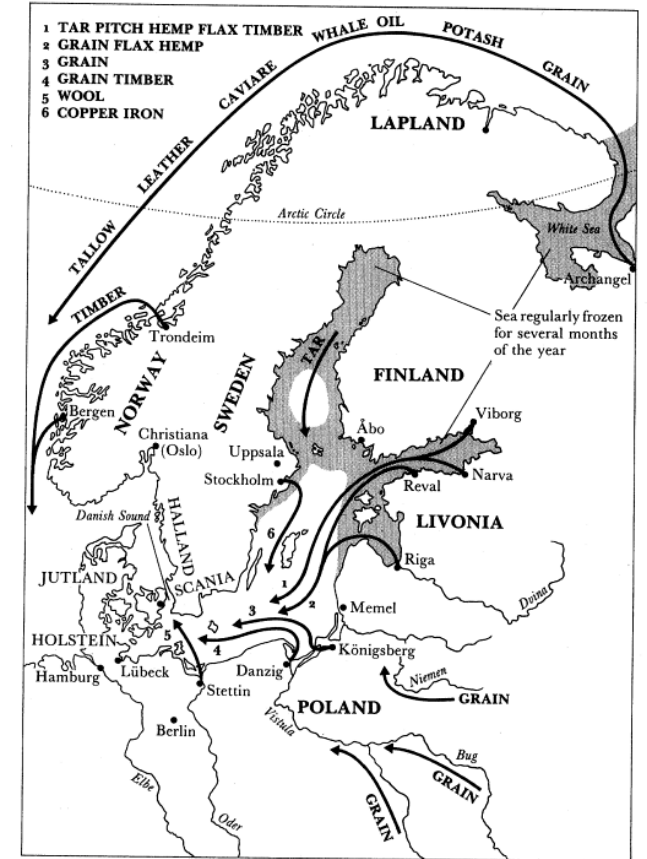
Linked Data



Linking Open Data cloud diagram 2014,
 by Max Schmachtenberg, Christian Bizer,
 Anja Jentzsch and Richard Cyganiak.
<http://lod-cloud.net/>

Dutch Ships and Sailors

- Dutch Naval history
- Documentation examples
- Linked Data for Dutch Ships and Sailors



Israel, 1989
<http://dutchshipsandsailors.nl>
de Boer et al., 2014

Table of Contents

- Introduction
- **Research Questions**
- Approach and Methodology
- Evaluation
- Recommendations
- Conclusion and Discussion

Research Questions

- How can additional datasets effectively be linked to those of the Dutch Ships and Sailors project?
- How can the additional datasets assist in answering existing research questions in the field of History?

Table of Contents

- Introduction
- Research Questions
- **Approach and Methodology**
- Evaluation
- Recommendations
- Conclusion and Discussion

Approach and Methodology

- General approach
- Datasets
- Requirements engineering
- Data conversion
- Visualisations



General approach

- Work for historians
- Evaluate each step, stay on track
- What do we need to answer the research questions?

Datasets

- Elbing and Archangel

1620 - #3

Op 14-4 heeft schipper Johan Peters uit Buiksloot met zijn schip, de Fortuna.

--»naar Elbing gebracht:				Tol per Product			Waarde per Product	
HG	4.0	Last	Haring					
HG	0.0	None	Kramerijen				300.0 fl.	
HG	102.0	Ohm	Rijnwijn					
HG	7.0	Pijp	Wijn					
Import Tol Totaal:								
«--vanuit Elbing naar Amsterdam meegenomen:				Tol per Product			Waarde per Product	
BF	4.0	Last	Rogge					
EG	20.0	Last	Gerst	7.0 m.	13.0 gr.	6.0 d.		
EG	42.54	Centner	Koper	14.0 m.	17.0 gr.	9.0 d.		
EG	20.0	Last	Tarwe	15.0 m.				
EG	88.0	Steen	Wol	11.0 m.				
HG	31.0	Last	Rogge					
HG	10.0	Steen	Veren					
Export Tol Totaal:				47.0 m. 30.0 gr. 15.0 d.				

Datasets

- Elbing and Archangel

Details van reis

Nummer	3325
Registratiedatum	1620-04-15
Bronverwijzing	SAA NA 162/88
Type akte	bevrachting
Bevrachters	Engelgraeff, Robert Valckenburch, Margr,Wed Vogelaer
Schipper	Govertsen, Willem
Herkomst schipper	Amsterdam
Naam schip	Engel
Naam schip (oorspronkelijk)	De Engel
Lastage	50.0
Gebruikte last	Russisch
Haven van vertrek	Amsterdam
Bestemmingshavens	Lapland Archangel Lapland Amsterdam
Vrachtprijs (totaal)	1670.0
Nummer De Buck	378

Requirements Engineering

- What is Requirements Engineering?
- Methodology chosen
- Results and Evaluation

Robson, 2011
Ebert, 2011

Results of Requirements Engineering

	Topic	Respondent 1	Respondent 2
Archangel	What makes Archangel interesting	Dataset isn't linked to DSS yet	Luxury products were traded for the entire golden age
	Exciting research Application examples	Research that offers new insights, a new take on matters. Economic growth compared to shipping quantity Ships that appear in both datasets Captains in both datasets Easily search for a dataset sample	Original research Geographical information
Elbing	What makes Elbing interesting	Data on load carried Good representation of baltic trade	Data on load carried
Archangel and Elbing	Research Questions	How big was shipping on Elbing/Archangel in total Dutch shipping How did wars influence shipping* How can climate/weather be linked to shipping Search for specific information in large datasets, such as certain names Price of goods, maybe over time. Do captains appear in multiple datasets* Can economic growth be linked to shipping* Do ships appear in multiple datasets*	Compare shipping to seasons Compare load data to DSS to see transfer of goods How did wars influence shipping* European shipping was usually only Dutch captains, considered more priviliged Traders who load ships can be compared, maybe see how finances were organised Baltic trade was less profitable but lower risk than other shipping, compare this Can economic growth be linked to shipping* Do captains appear in multiple datasets* Do ships appear in multiple datasets*

Results of Requirements Engineering

Exhaustive list of research topics or questions	How to represent after conversion
How big was shipping on Elbing/Archangel in total Dutch shipping	Compare datasets to DSS
How did wars influence shipping*	War times from external data, compare to volumes of shipping
How can climate/weather be linked to shipping	Weather data, compare to shipping volumes
Search for specific information in large datasets, such as certain names	Can be resolved with Query
How do the price of goods change over time	Can be resolved with Query
Do captains appear in multiple datasets*	Compare datasets to DSS
Can economic growth be linked to shipping*	Compare shipping volumes to economic data from external resource
Do ships appear in multiple datasets*	Can be resolved with Query
Compare load data to DSS to see transfer of goods	Query and compare to DSS
European shipping was usually only Dutch captains, considered more privilliged	Query to verify, compare to DSS
Traders who load ships can be compared, maybe see how finances were organised	Does not appear in both dataset
Baltic trade was less profitable but lower risk than other shipping, compare this	Query and compare to DSS

Data conversions

- Current data organisation
- Design choices
- Existing tools or start over?
- Quick conversion vs. Specific conversion

Data conversions

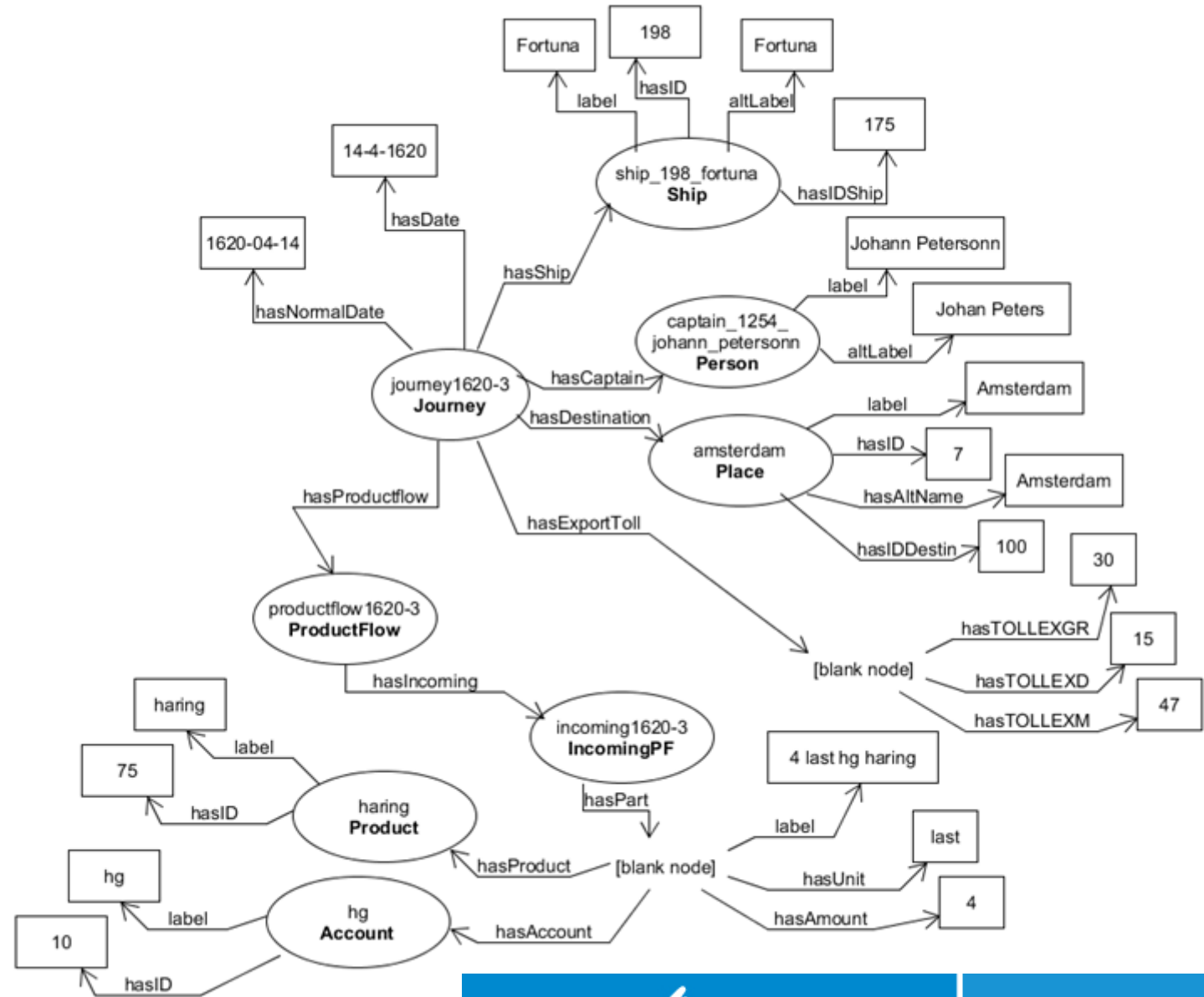
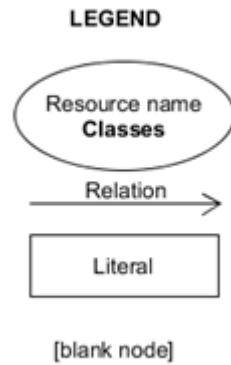
```
<table name="journey">
  <column name="YEAR">1620</column>
  <column name="NR">3</column>
  <column name="DAY">14</column>
  <column name="MONTH">4</column>
  <column name="SHIP_ID">198</column>
  <column name="CAPTAIN_ID">1254</column>
  <column name="DESTIN_ID">7</column>
  <column name="TOLLIMFL">0</column>
  <column name="TOLLIMM">0</column>
  <column name="TOLLIMGR">0</column>
  <column name="TOLLIMSCH">0</column>
  <column name="TOLLIMD">0</column>
  <column name="TOLLEXFL">0</column>
  <column name="TOLLEXM">47</column>
  <column name="TOLLEXGR">30</column>
  <column name="TOLLEXSCH">0</column>
  <column name="TOLLEXD">15</column>
</table>
```

```
<table name="voyage">
  <column name="id">8346</column>
  <column name="number">378</column>
  <column name="date">1620-04-15</column>
  <column name="source">SAA NA 162/88</column>
  <column name="source_type">bevrachting </column>
  <column name="freighter1">Engelgraeff, Robert</column>
  <column name="freighter2">Valckenburch, Margr,Wed Vogelaer</column>
  <column name="freighter3"></column>
  <column name="freighter4"></column>
  <column name="freighter5"></column>
  <column name="captain">Govertsen, Willem </column>
  <column name="captain_from">Amsterdam</column>
  <column name="ship">De Engel</column>
```

Data conversions

- Current data organisation
- Design choices
- Existing tools or start over?
- Quick conversion vs. Specific conversion

Data conversions



Data conversions

- Current data organisation
- Design choices
- Existing tools or start over?
- Quick conversion vs. Specific conversion

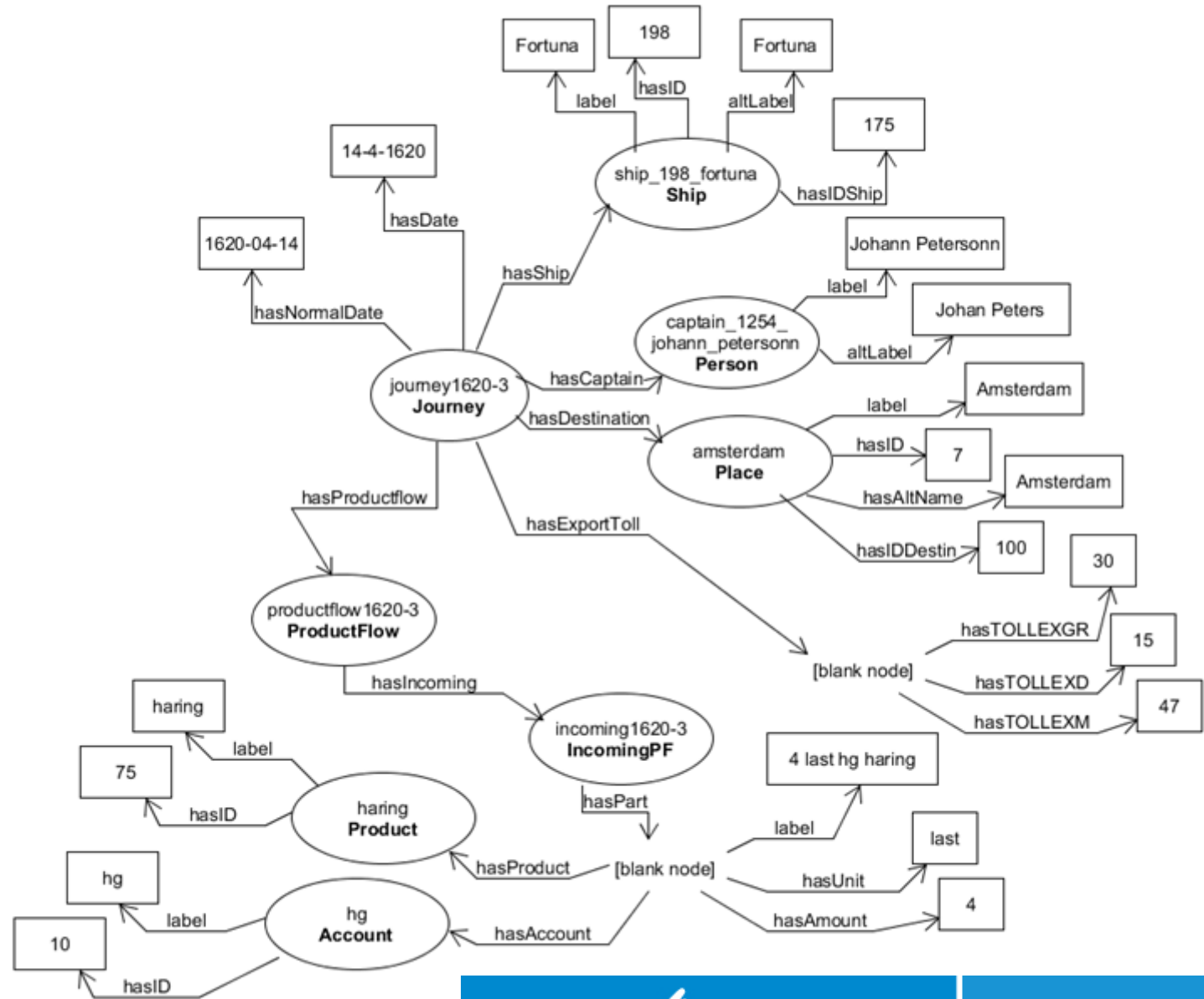
Data conversions

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix elb: <http://purl.org/collections/nl/dss/elbing/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
elb:journey1620-3
  a elb:Journey ;
  elb:hasDate "14-4-1620" ;
  elb:hasNormalDate "1620-04-14"^^xsd:date ;
  elb:hasShip elb:ship_198_fortuna ;
  elb:hasCaptain elb:captain_1254_johann_petersonn ;
  elb:hasDestination elb:amsterdam ;
  elb:hasExportToll
    [a elb:Toll ;
     elb:hasTOLLEXM "47" ;
     elb:hasTOLLEXGR "30" ;
     elb:hasTOLLEXD "15" ] ;
  elb:hasProductflow elb:productflow1620-3 .
```

```
elb:ship_198_fortuna
  a elb:Ship ;
  elb:label "Fortuna" ;
  elb:hasID "198" ;
  elb:AltLabel "Fortuna" ;
  elb:hasIDShip "175" .
```

```
elb:captain_1254_johann_petersonn
  a elb:Person ;
  elb:hasID "1254" ;
  elb:hasIDCaptain"880" ;
  elb:hasHometown elb:buiksloot ;
  elb:label "Johann Petersonn" ;
  elb:altLabel "Johan Peters" .
```



Data conversion (cont.)

- Mapping data to Dutch Ships and Sailors
- Conversion results
- Noteworthy remarks on Elbing
- Noteworthy remarks on Archangel

Data conversion (cont.)

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix elbing: <http://purl.org/collections/nl/dss/elbing/> .

elbing:Journey
  a rdfs:Class ;
  rdfs:label "Journey" .

elbing:Person
  a rdfs:Class ;
  rdfs:label "Person" .

elbing:Place
  a rdfs:Class ;
  rdfs:label "Place" .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ns2: <http://purl.org/collections/nl/dss/elbing/> .
@prefix dss: <http://purl.org/collections/nl/dss/> .
@prefix mdb: <http://purl.org/collections/nl/dss/mdb/> .
@prefix das: <http://purl.org/collections/nl/dss/das/> .
@prefix archangel: <http://purl.org/collections/nl/dss/archangel/> .

ns2:Journey
  a rdfs:Class ;
  rdfs:subClassOf dss:Record, dss:Voyage ;
  rdfs:label "Journey" .

ns2:Person
  a rdfs:Class ;
  rdfs:subClassOf foaf:Person, dss:Person;
  rdfs:label "Person" .

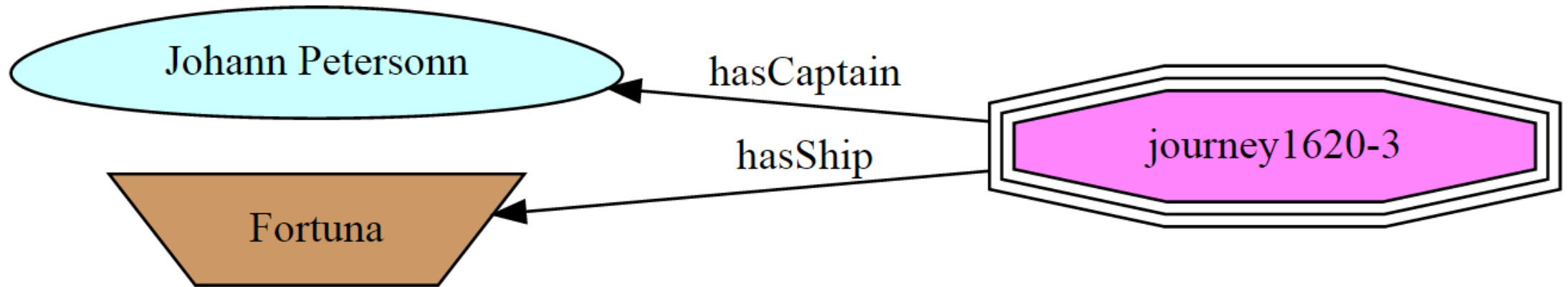
ns2:Place
  a rdfs:Class ;
  rdfs:subClassOf dss:Place, skos:Concept;
  rdfs:label "Place" .
```

Data conversion (cont.)

- Mapping data to Dutch Ships and Sailors
- Conversion results
- Noteworthy remarks on Elbing
- Noteworthy remarks on Archangel

Data conversion (cont.)

Context graph



Data conversion (cont.)

- Mapping data to Dutch Ships and Sailors
- Conversion results
- Noteworthy remarks on Elbing
- Noteworthy remarks on Archangel

Visualisations

- Design based on research questions from requirements engineering
- Proof of Concept visualisations
- GeoNames Data linked in Heatmap
- Interlinking of Data in a query, present in chart

Heatmap

```
PREFIX arch: <http://purl.org/collections/nl/dss/archangel/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX elb: <http://purl.org/collections/nl/dss/elbing/>
PREFIX wgs: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX gn: <http://www.geonames.org/ontology#>
```

```
SELECT (COUNT(?name) AS ?count) ?name ?lat ?long WHERE {
{
SELECT ?name ?lat ?long
WHERE      {
?journey elb:hasDestination ?stad.
?stad owl:sameAs ?placegeo.
?placegeo wgs:lat ?lat .
?placegeo wgs:long ?long .
?placegeo gn:name ?name .
}
} UNION {
SELECT ?name ?lat ?long
WHERE      {
?journey arch:hasHarbour1 ?stad.
?stad owl:sameAs ?placegeo.
?placegeo wgs:lat ?lat .
?placegeo wgs:long ?long .
?placegeo gn:name ?name .
}
}
} GROUP BY ?name ?lat ?long"
```



Visualisations

- Design based on research questions from requirements engineering
- Proof of Concept visualisations
- GeoNames Data linked in Heatmap
- Interlinking of Data in a query, present in chart

Chart

PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX elb: <http://purl.org/collections/nl/dss/elbing/>
PREFIX arch: <http://purl.org/collections/nl/dss/archangel/>

```
SELECT ?year (COUNT(?elb + ?arch) AS ?count) (COUNT(?arch) AS ?carch) (COUNT(?elb) AS ?celb) WHERE {  
{  
SELECT * WHERE {  
  ?elb elb:hasNormalDate ?date .  
}  
}  
} UNION {  
  
SELECT * WHERE {  
  ?arch arch:hasNormalDate ?date .  
}  
}  
}  
GROUP BY (year(?date) as ?year)
```

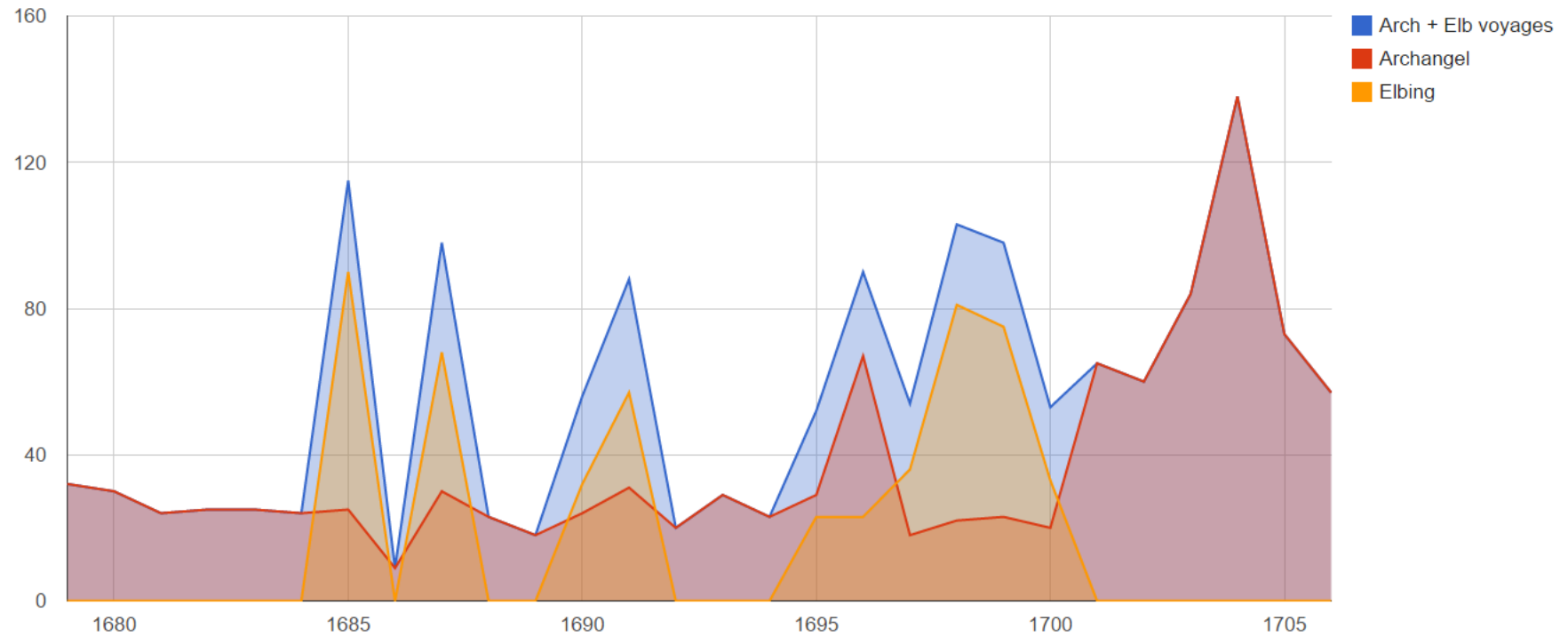


Table of Contents

- Introduction
- Research Questions
- Approach and Methodology
- **Evaluation**
- Recommendations
- Conclusion and Discussion

Evaluation

- Evaluation questions
- Results of the evaluation
 - Strange spikes, underlying data at fault?
 - Relatively little effort to create powerful tool
 - Research questions can not be answered, not enough data

Table of Contents

- Introduction
- Research Questions
- Approach and Methodology
- Evaluation
- **Recommendations**
- Conclusion and Discussion

Recommendations

1. Set up data requirements
 - Try to stay close to original data
2. Predefine RDF structure based on requirements
3. Convert tables based on RDF structure
4. Map to existing concepts for data enrichment

Table of Contents

- Introduction
- Research Questions
- Approach and Methodology
- Evaluation
- Recommendations
- **Conclusion and Discussion**

Discussion and Conclusion

- Data problems
- Visualisation problems
- More historic data to Linked Data
- Research questions answered

Discussion and Conclusion

- How can additional datasets effectively be linked to those of the Dutch Ships and Sailors project?
 - Four step recommendations
 - Converting to Linked Data is possible without loss of information
- How can the additional datasets assist in answering existing research questions in the field of History?
 - Data is enriched by creating a bigger picture
 - Not enough data in Dutch Ships and Sailors yet

Future Work

- More data needed!
- Date standardisation
- Some aspects of data left untouched