

Disclosing Civil Servants Registries as Linked Open Data

Richard L. Zijdeman (IISG, University of Stirling)
Cambridge Social Stratification Seminar
Sep 7-8, 2022, Stirling University, Stirling, UK

Problem 1: representation of women's work in historical sources

- Underregistration of work activities
- Unrecognized activities
- Only allowed in certain occupations

Exception: domestic servant registers

- Registration of mainly women
- Available for multiple decades
- Available in multiple countries
- Detailed info on age, religion and migration
- (But indeed limited to a number of occupations)

Been there, done that, bought the T-shirt?

- Domestic service work as ‘bridging’ occupation;
- Move to urban areas to ‘escape’ rural lower positions;
- Acquire knowledge on ways of living of moderate and higher classes;
- Gain income for parental household (and reducing cost of extra child);

E.g.: Bras 2003, McBride 1974, Laslett 1965, Fauve-Chamoux & Wall 2012

Problem 2: representation of local archives

- Local archives do most of the hard work:
 - preserving data, digitizing data, data entry
 - Get no credit for any of it from ‘us’ (not even ‘clicks’)
 - Making it more difficult to secure funding
 - (Also bad for ‘us’)

https://www.ted.com/talks/tim_berners_lee_on_the_next_web

The Next Web





Amsterdam

Capital of the Netherlands

Amsterdam is the Netherlands' capital, known for its artistic heritage, elaborate canal system and narrow houses with gabled facades, legacies of the city's 17th-century Golden Age. Its Museum District houses the Van Gogh Museum, works by Rembrandt and Vermeer at the Rijksmuseum, and modern art at the Stedelijk. Cycling is key to the city's character, and there are numerous bike paths.

Area: 219.3 km²

Weather: 1°C, Wind NE at 18 km/h, 80% Humidity

Local time: Thursday 21:29

Population: 821,752 (2015) UNdata

Plan a trip

[Amsterdam travel guide](#)

[3-star hotel averaging €111, 5-star averaging €283](#)

[Upcoming Events](#)

Colleges and Universities: [University of Amsterdam](#), [MORE](#)

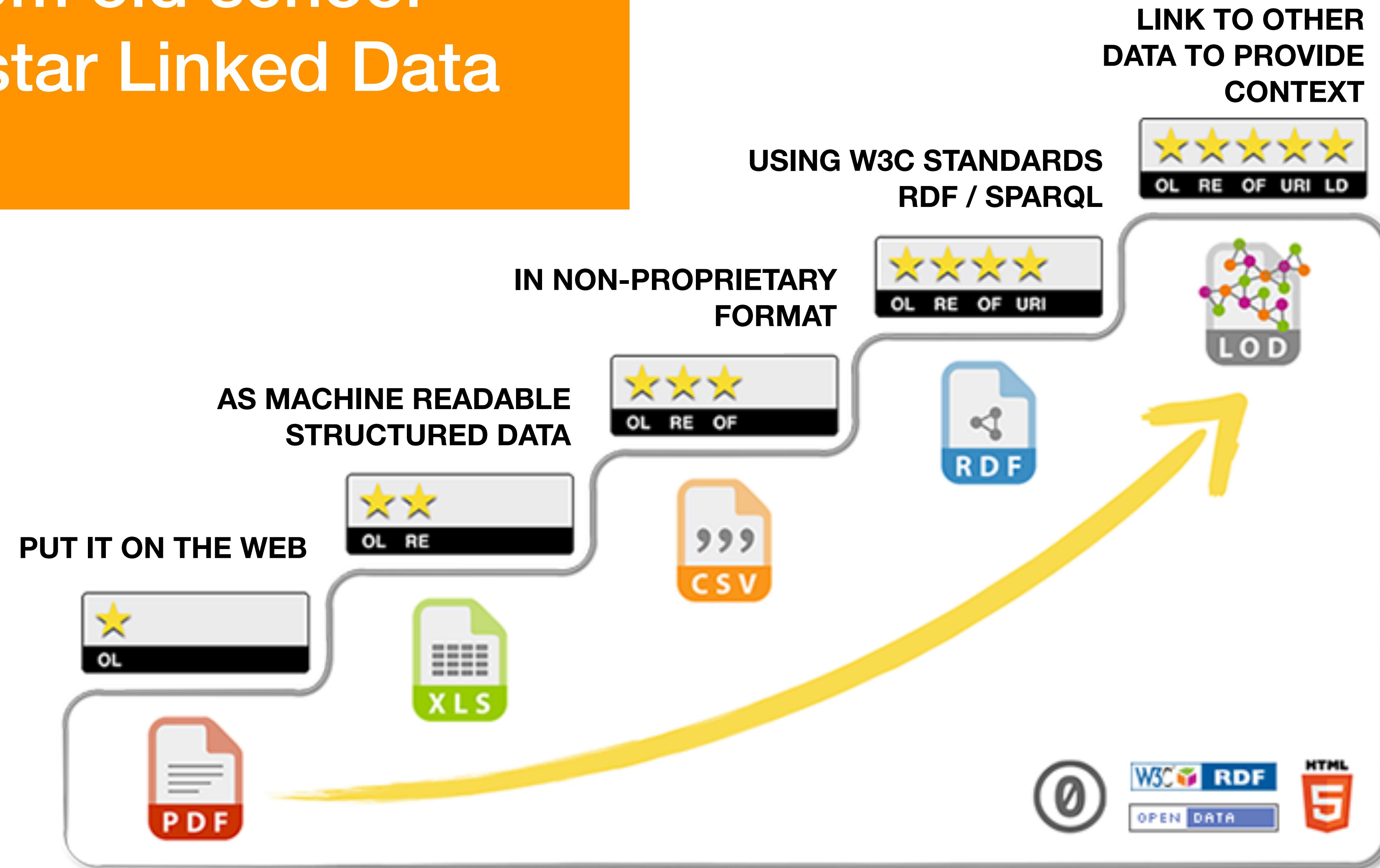
People also search for



You've
probably
used it
already!

Harderwijk	
Municipality	
	Harbour of Harderwijk seen from windmill De Hoop
	Location in Gelderland
	Coordinates: 52°21'N 5°37'E
Country	Netherlands
Province	Gelderland
Government ^[1]	
• Body	Municipal council
• Mayor	Harm-Jan van Schaik (CDA)
Area ^[2]	
• Total	48.27 km ² (18.64 sq mi)

From old school to 5-star Linked Data



Linked Data (RDF)

Represent data facts as URI's:

- Unique ID, that's preferably also a URL (retrievable via the Web)
- “Bridge of Allan” -> [https://dbpedia.org/page/Bridge of Allan](https://dbpedia.org/page/Bridge_of_Allan)

Linked Data (RDF)

Structure facts in triples (tripartite statements):

ID	birthPlace
1	Bridge of Allan
...	...

<id1>

<hasBirthPlace>

<bridgeOfAllen>

<<https://mydata.org/id/1>>

<<https://schema.org/birthPlace>>

<[https://dbpedia.org/page/Bridge of Allan](https://dbpedia.org/page/Bridge_of_Allan)>

Data

- Dienstboden ('domestic service') register 1888-1909
 - [1873-1880, 1880-1887, 1906-1919]
- Data entry by volunteers of the Historical Association 'Herderwijk'
- Data provided by Historical Knowledge Center Harderwijk
 - Basic cleaning (mainly of dates and incorrect fields)
 - N_unique = 1519 (N_raw=1880)



Variables

Dienstbodenregister Harderwijk 1888-1909

inventaris NAH 940

inschr	3=familienaam	4=voornaam	geb.dd	geb.pl.	9=BS 11=religie	12=beroep	13=huizing	14=aan komst	15=vorige woonplaats	16=vertrek	17=naar	20=opmerkingen
---	Apperlo	Elisabeth	10-11-1864	Zwartsluis	o NH	dienstmeid	A. Bilderbak de Hondt	17-03-1884	Amsterdam	24-09-1887	Zwartsluis	Blad 1
---	---	---	---	---	---	---	A. Willemsen	---	---	---	---	---
---	Appelboom	Jannetje	24-05-1868	Ermelo	o NH	dienstmeid	H. Kamm	17-10-1884	Nijkerk	---	Deel D blz. 96	Gehuwd 06-02-1895
---	---	---	---	---	---	---	Ephraim	---	---	---	---	---
---	---	---	---	---	---	---	Melms	---	---	---	---	---
---	Aartsen	Johanna Everdina	21-09-1864	Harderwijk	o NH	dienstmeid	P.H. Thijs	---	---	---	Deel A blz. 6	---
---	---	---	---	---	---	---	Garjeanne	---	---	---	---	---
---	Aartsen	Jans	02-07-1861	Harderwijk	o NH	dienstmeid	Jonkheer	---	---	02-05-1886	Hoornaar	---

- familyName
- (civil status)
- hostName
- placeComingFrom
- givenName
- religion
- comments
- dateOfArrival
- dateOfBirth
- (occupation)
- placeLeavingTo
- placeOfBirth
- dateOfDeparture

Data conversion

- Data converted using a ; tool: CoW (Meroño-Peñuela et al. 2019)
- Takes a csv as argument
- Automatically creates RDF and a ‘recipe’ for the conversion
- You can manually alter the recipe (like an R/Python script)
- Recipe available via: <https://github.com/rlijdeman/hkh-maids>

Data ‘live’ available

- Hosted via Triply’s triple store:

<https://triplydb.com/rlzijdeman/hkh-dienstboden-register>

The screenshot shows the Triply triple store interface. At the top, there's a blue header bar with the URL 'https://triplydb.com/rlzijdeman/hkh-dienstboden-register'. On the left, a sidebar menu includes 'Dienstboden Regist...', 'Browser', 'Table', 'SPARQL', 'Graphs', 'Services', 'Assets', and 'Insights'. The main content area displays the dataset details: 'Dienstboden Register Harderwijk 1888-1909' by 'Silk', created a year ago with 37.601 statements. It includes a thumbnail image of a building, a description of the register of maids ('dienstboden') of Harderwijk, and links to queries and graphs. Below this, there are sections for 'Example resources' (with links to 1866, 569, 1140, 569) and 'Dependent queries' (with links to hkh-average-stay-by-year, servants-stay-duration-1, and servants-stay-duration). A 'Search ...' bar is at the top right.

< 569

Person

<http://www.hkharderwijk.nl/nah/NAH940-DBRegister-1888-1909/persons/servant/569>

BirthDate

1863-01-04

BirthPlace

Harderwijk

FamilyName

Haverkamp

GivenName

Aaltje

HasOccupation

dienstbode

CivilStatus

o

Religion

NH

Type

Person

API

GET: https://api.triplydb.com/queries/rizijdeman/hkh-average-stay-by-monthh/run

</>

Variables +

```
1 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
2 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX nah: <http://nah.org/ns/>
5
6 #SELECT ?dep
7
8 #SELECT ?dept
9
10 #SELECT ?year
11
12 #SELECT ?dept
13
14 #SELECT ?dept
15
16
17
18
19
20
21
22
23
24
25
26
27
```

Code snippets

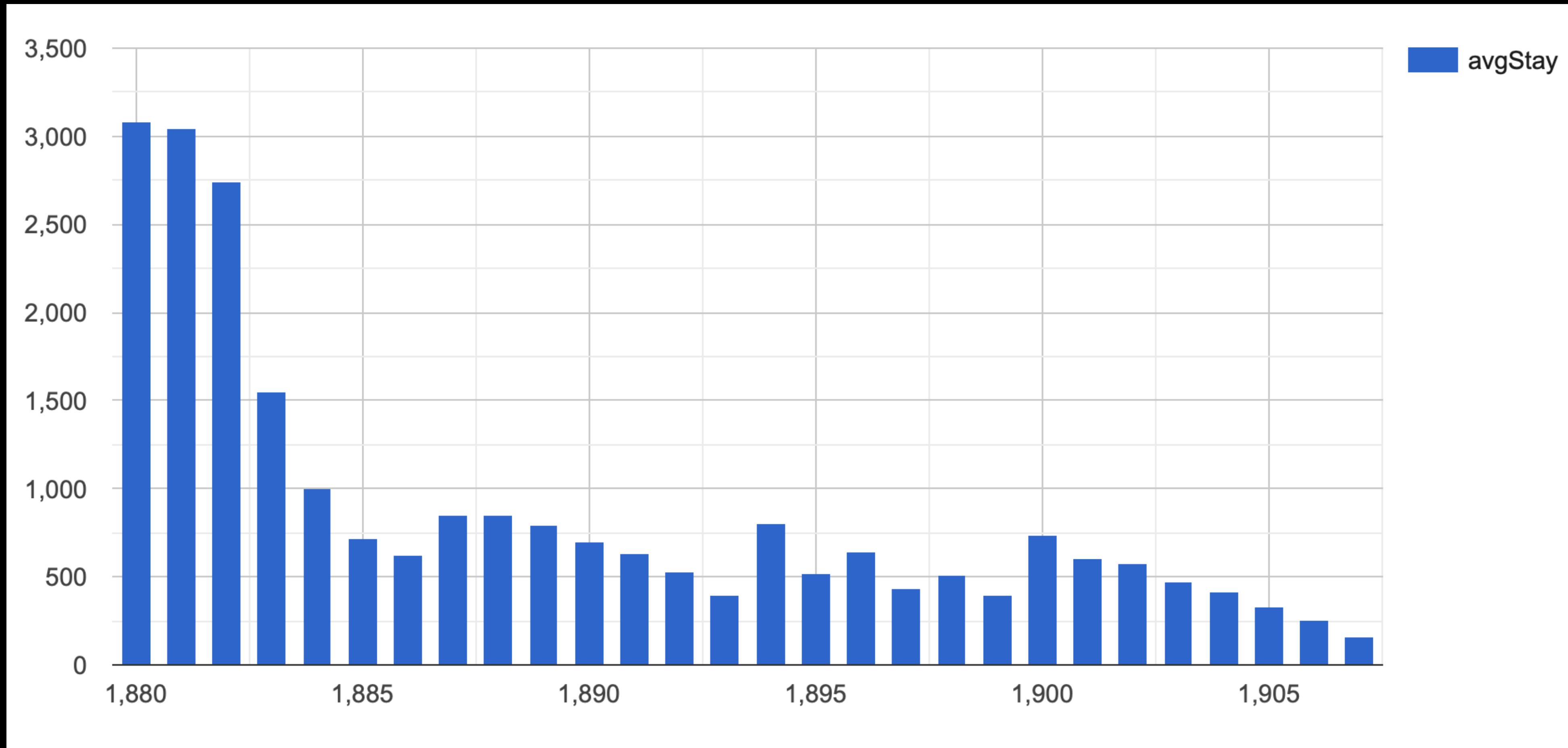
PYTHON

R

```
library(httr) # for getting data
library(jsonlite) # for working with JSON
URL = "https://api.triplydb.com/queries/rizijdeman/hkh-average-stay-by-monthh/run?pageSize=100"
data <- GET(URL)
data <- fromJSON(content(data, as = 'text', encoding = "UTF-8"))
```

COPY TO CLIPBOARDCLOSE

Average stay in days



Where were maids born?

1	"545" ^{^^xsd:integer}	"Harderwijk"@nl
2	"42" ^{^^xsd:integer}	"Amsterdam"@nl
3	"28" ^{^^xsd:integer}	"Zwolle"@nl
4	"13" ^{^^xsd:integer}	"Amersfoort"@nl
5	"13" ^{^^xsd:integer}	"Haarlem"@nl
6	"12" ^{^^xsd:integer}	"Meppel"@nl
7	"11" ^{^^xsd:integer}	"Zwartsluis"@nl
8	"9" ^{^^xsd:integer}	"Arnhem"@nl
9	"9" ^{^^xsd:integer}	"Assen"@nl
10	"7" ^{^^xsd:integer}	"Doornspijk"@nl
11	"6" ^{^^xsd:integer}	"Hattem"@nl
12	"5" ^{^^xsd:integer}	"Dordrecht"@nl
13	"5" ^{^^xsd:integer}	"Hilversum"@nl
14	"4" ^{^^xsd:integer}	"Leiden"@nl
15	"4" ^{^^xsd:integer}	"Steenwijk"@nl
16	"4" ^{^^xsd:integer}	"Tilburg"@nl
17	"4" ^{^^xsd:integer}	"Weerselo"@nl

- Places are provided as strings
- Most places are available via DBpedia
- So we can query DBpedia for coordinates...
- and depict the results

Where were maids born?

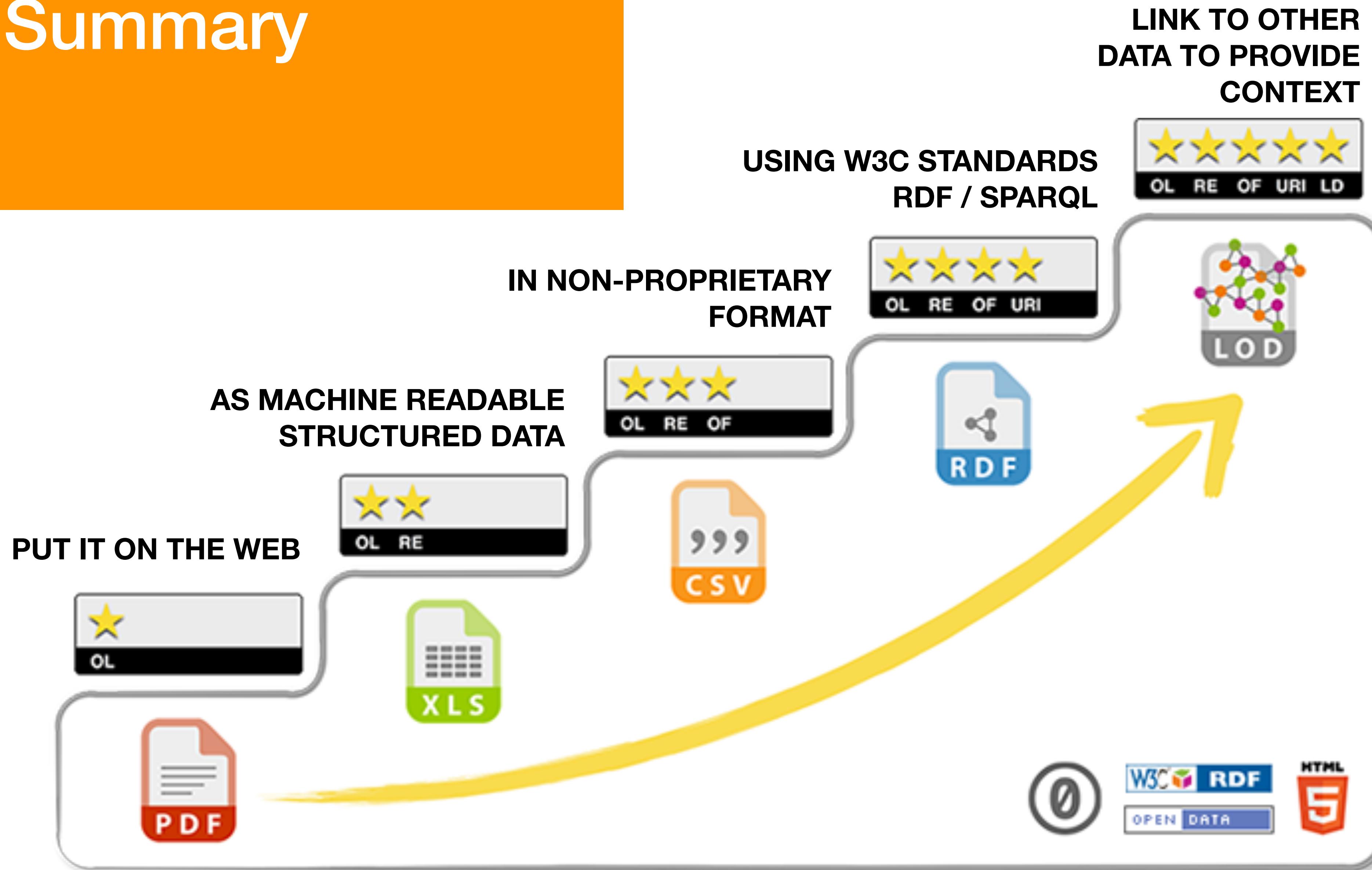
```
1 PREFIX sdo: <https://schema.org/>
2 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX dbo: <http://dbpedia.org/ontology/>
5 PREFIX georss: <http://www.georss.org/georss/>
6 PREFIX geo: <http://www.opengis.net/ont/geosparql#>
7 PREFIX dbr: <http://dbpedia.org/resource/>
8
9 # Thanks to Wouter Beek for helping me out with this query!
10
11 select ?count ?birthPlace ?coordinate ?coordinateLabel {
12 {
13   select (strlang(?birthPlaceString,'nl') as ?birthPlace) (count(?sub) as ?count) {
14     ?sub sdo:birthPlace ?birthPlaceString. # all individuals with a birthplace
15   }
16   group by ?birthPlaceString
17   order by desc(?count)
18 }
19 # retrieve for all places in the Netherlands the coordinates
20 service <https://dbpedia.org/sparql> {
21   [] a dbo:Place;
22   dbo:country dbr:Netherlands; # to limit the number of results, we don't want all places in the world
23   rdfs:label ?birthPlace;
24   georss:point ?Point .
25   # DBPEDIA seems to have a different ordering of lat long so turning that around
26   BIND(CONCAT(STRAFTER(?Point, " "), " ", STRBEFORE(?Point, " ")) as ?reversePoint)
27   # turning our coordinates literal into a proper geo-type
28   BIND(STRDT(CONCAT('POINT(',STR(?reversePoint),)'),geo:wktLiteral) as ?coordinate)
29   #FILTER ( str( ?wktLabel) = "Ermelo")
30 }
31 bind('''<h3>{{birthPlace}}</h3><p>Count: {{count}}</p>'''^^rdf:HTML as ?coordinateLabel)
32 }
33 limit 17 # 4 or more occurrences
```



Query at:

<https://triplydb.com/rlzijdeman/-/queries/Query-1/5>

Summary



Next steps

- Advocate this approach to the NDE (Dutch National Heritage Network)
- Enhance sample with (linked) data:
 - Use burgerLinker to link data to LINKS (civil registers)
 - Expand to other (Harderwijk) local samples
 - Also convert other variables (such as religion and occupation)