



CLARIAH – a Dutch distributed research infrastructure for the humanities and social sciences

Data for History workshop, Lyon, 24-25 May 2018

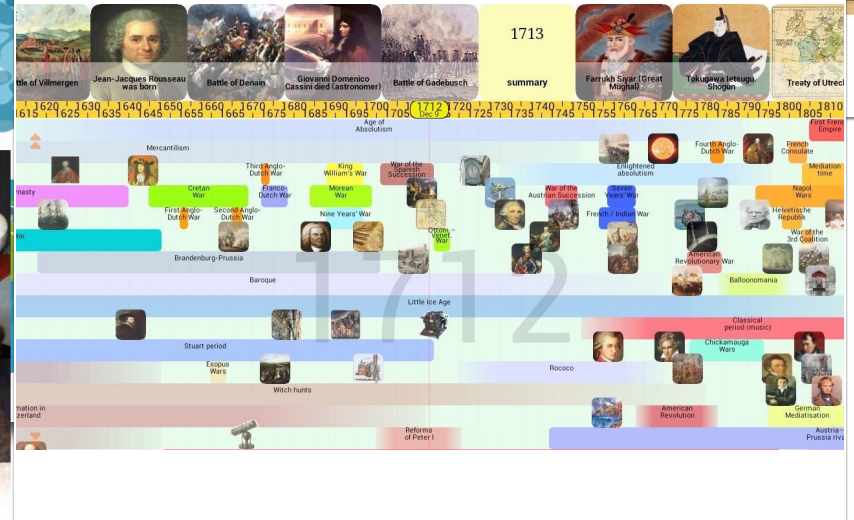
Lodewijk Petram
lodewijk.petram@huygens.knaw.nl



Concept

Infrastructure

Data modelling



Concept

Infrastructure

Data modelling

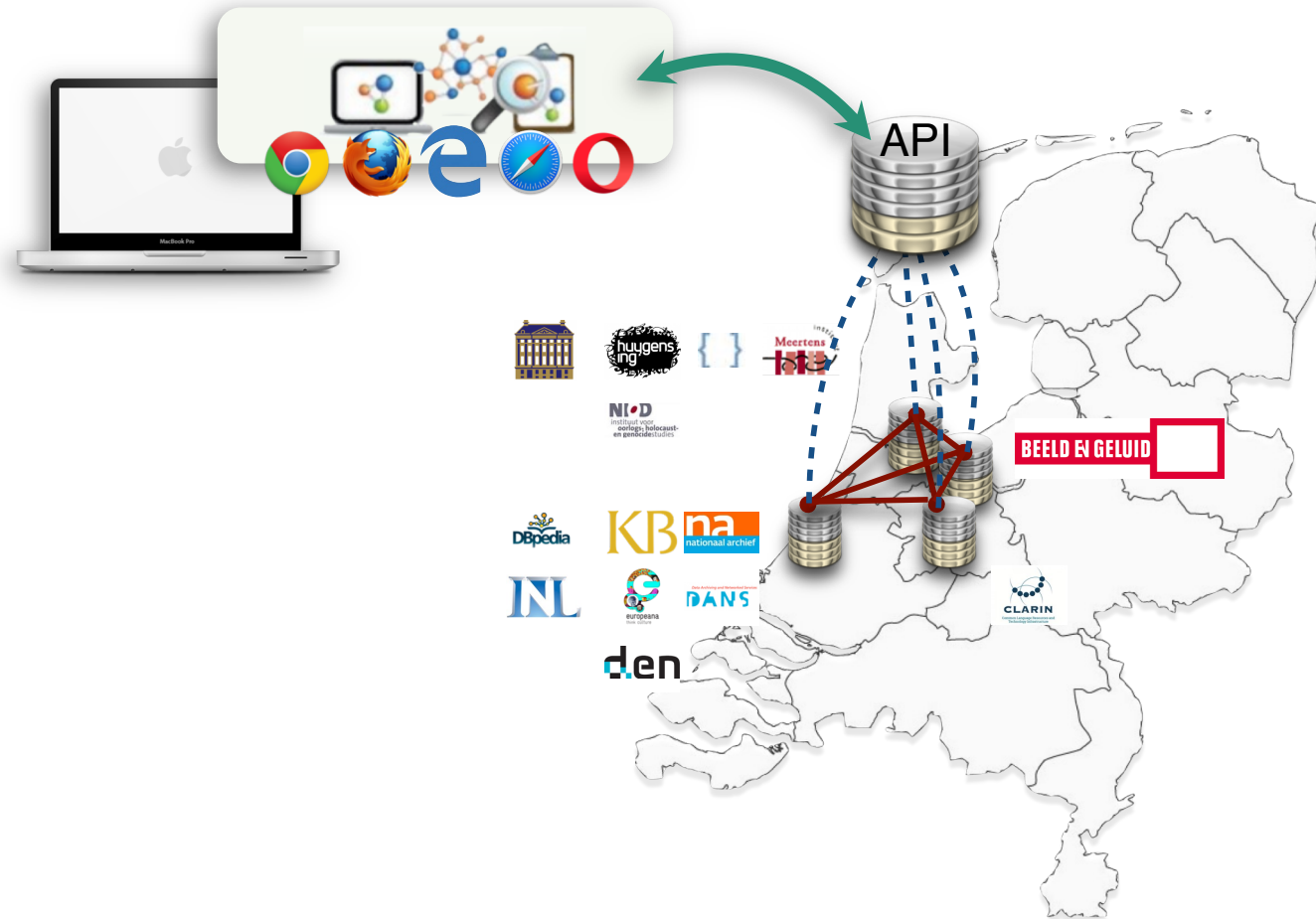


A screenshot of a web browser displaying the ANANSI website. The browser's address bar shows the URL 'https://anansi.clariah.nl/?hsid=0958417f-9433-4b09-8e47-03643b5ead13'. The website header includes the CLARIAH logo and a user profile for 'Lodewijk'. The main content area features a large banner with the title 'Bridge to networked research data' and a descriptive paragraph: 'ANANSI is the central data connection between the three primary CLARIAH domains (Linguistics, Social & Economical History and Media Studies). ANANSI facilitates the connectivity of structured (meta-)data served by the research infrastructures constructed in these domains. Furthermore, ANANSI integrates data with large scale existing data infrastructures outside CLARIAH - both within and outside the humanities.' Below the banner, a section titled 'Featured datasets' displays three categories: 'Persons' (a grid of portrait photos), 'Concepts' (a document with a complex diagram), and 'Locations' (a map with highlighted areas). A vertical blue bar on the right side of the page reads 'Powered by' followed by the logo of the underlying infrastructure.

Concept

Infrastructure

Data modelling



Concept

Infrastructure

Data modelling

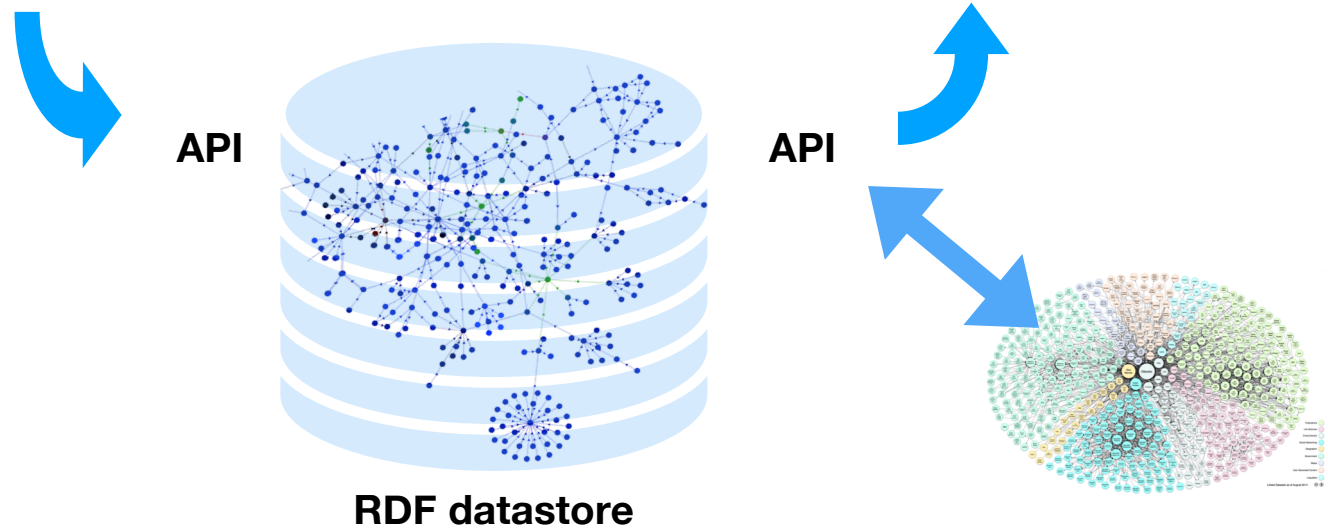


Upload of data

- ◆ supported formats: csv, excel, MS access, RDF
- ◆ remote repo's (ResourceSync)
- ◆ future integration with tools like OpenRefine

Use of data

- ◆ faceted search
- ◆ export in csv
- ◆ developer API on GraphQL
- ◆ future integration with tools like Gephi





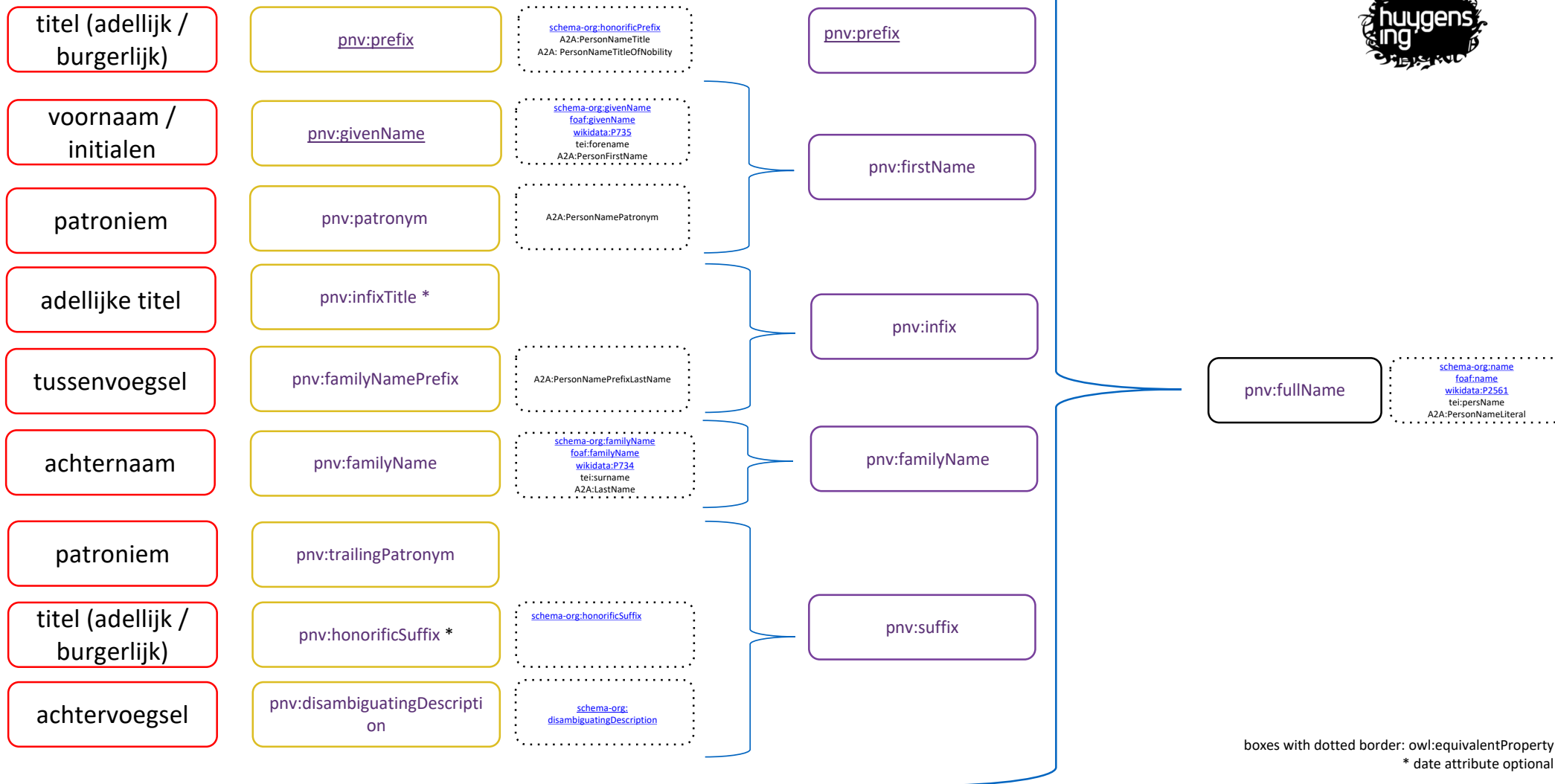
Concept

Infrastructure

Data modelling

- ◆ Software is schema agnostic; generates a schema on data import
- ◆ Mapping GUI (under development) allows users to create a schema that suits their needs
- ◆ Challenge: how to gently force users to use a set of standard ontologies?

Person Name Vocabulary — name elements

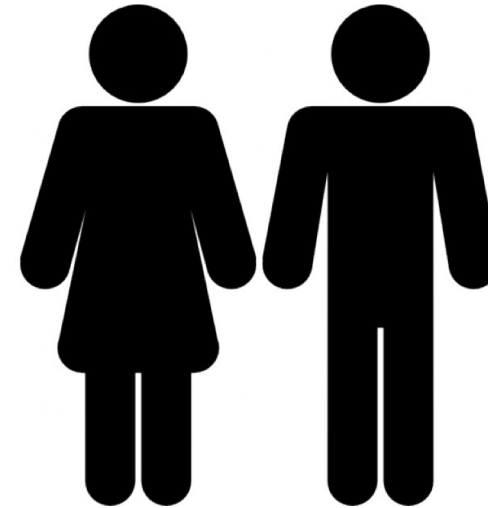


boxes with dotted border: owl:equivalentProperty
* date attribute optional

Additional identifying elements



volledige naam	pnv:fullName	schema-org:name foaf:name wikidata:P2561 tei:persName A2A:PersonNameLiteral
geslacht	schema-org:gender	foaf:gender wikidata:P21 A2A:gender
geboortedatum	schema-org:birthDate	wikidata:P569 A2A:BirthDate
geboorteplaats	schema-org:birthPlace	wikidata:P19 A2A:BirthPlace
sterfdatum	schema-org:deathDate	wikidata:P570 A2A:DeathDate
sterfplaats	schema-org:deathPlace	wikidata:P20 A2A:DeathPlace
Doopdatum + plaats	wikidata:P1636	
Begraafdatum + plaats	wikidata:4602	



[schema-org:person](#)

[cidoc:E21_person](#)
[foaf:person](#)
[wikidata:Q5](#)
(OR (fictional):
[wikidata:Q15632617](#))



Thank you!

<https://clariah.nl/>

<https://anansi.clariah.nl/>

<https://data.anansi.clariah.nl/>

lodewijk.petram@huygens.knaw.nl

