



Transformations for ~~aggregating~~ Linked Open Data

The Local Amsterdam Cultural Heritage
Linked Open Data Network

Lukas Koster (*Library of the University of Amsterdam*)
Ivo Zandhuis (*Ivo Zandhuis Research & Consultancy*)

SWIB 2018 Bonn



AdamLink Project

AdamNet Foundation:

Amsterdam based Library Collaboration Association

33 member institutions

- public libraries
- higher education libraries
- special libraries
- museums
- archives
- documentation centers



AdamLink Project Goal

Create a linked open data infrastructure

for AdamNet member institutions collections

on the topic of "*Amsterdam*"

Data Only!

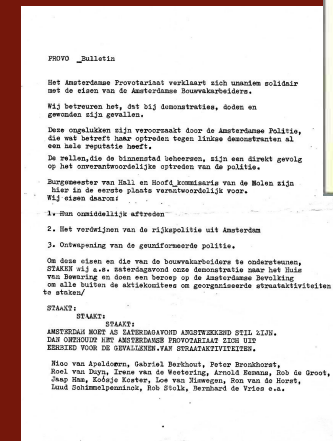
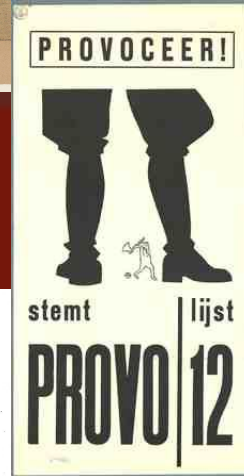
And some proof of
concept apps

Target audiences

- ≡ Researchers
- ≡ Education
- ≡ General public
- ≡ Creative industry



Initial Collections



Photographs,
paintings,
drawings,
posters, maps,
etc.

Amsterdam Museum

Image database

Amsterdam City Archive

Image database

International Institute of Social History

Image database

Interviews

Texts

University of Amsterdam

Digitised maps

Amsterdam Public Library

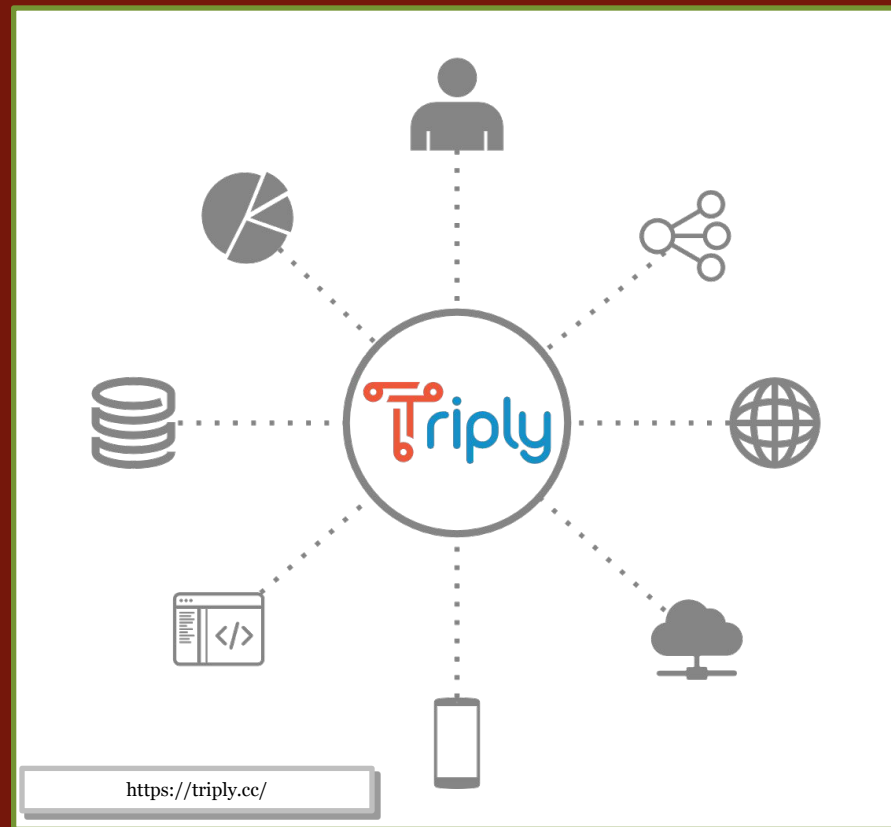
Books on topic "Amsterdam"

Catalogue records



Triply Linked Open Data Platform

- ≡ Triple Store
- ≡ Import Turtle
- ≡ Browser
- ≡ Tables
- ≡ Sparql Endpoints
- ≡ API
- ≡ Management
- ≡ Hosted



Challenge:

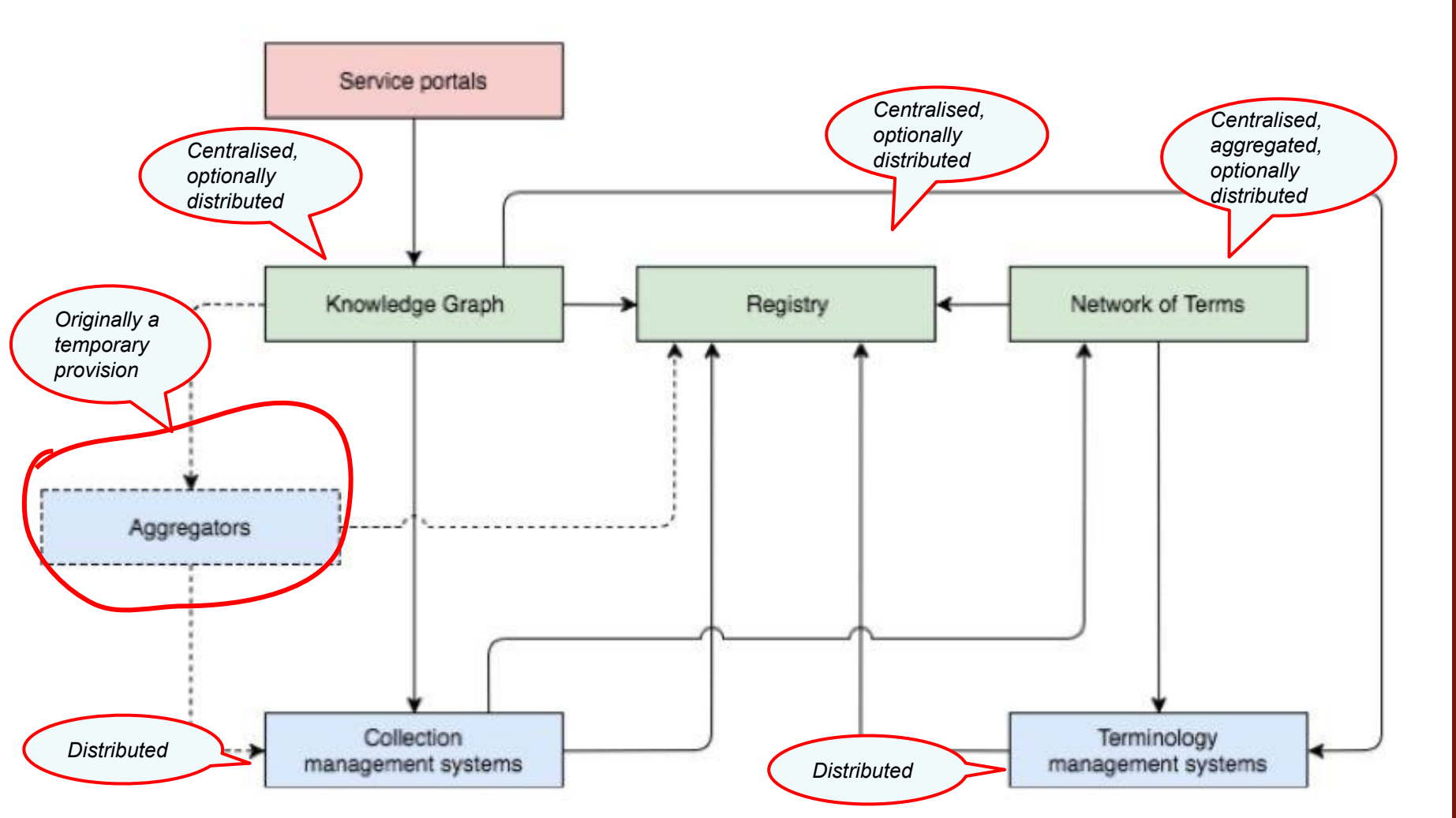
Distribution vs. Aggregation

Distribution is up-to-date, but

- ≡ Often: no 'live' LOD at all
- ≡ Local differences in ontologies, vocabularies, data types, data quality
- ≡ Complicated data joining
- ≡ Performance issues

Dutch National Program "NDE" Digital Heritage Network

<http://www.netwerkdigitaalervoed.nl/en>





Aggregators: changing role

NDE: extension to the model - “service platforms”

“A service platform combines and enriches heritage information and makes it usable in a specific context”

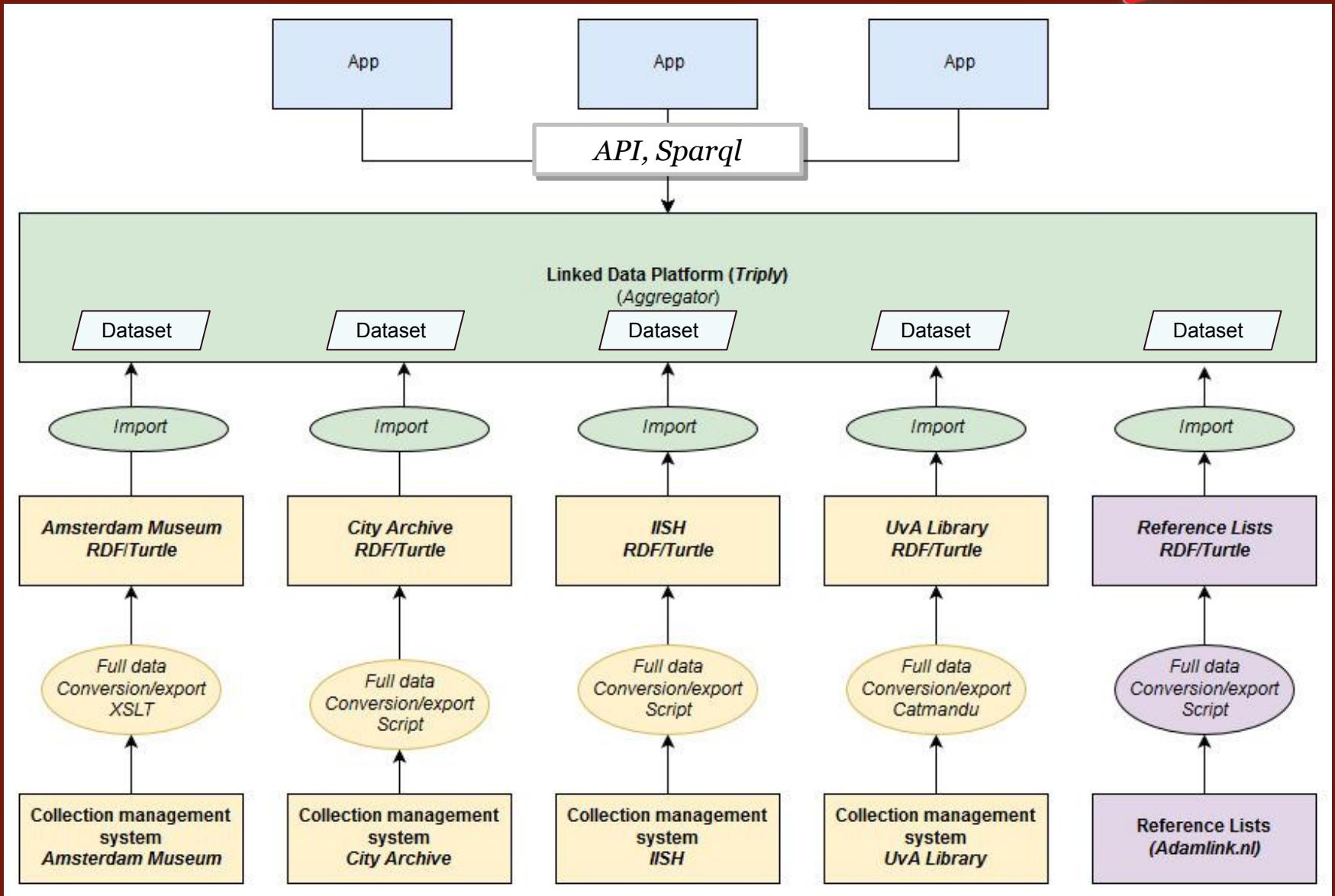
(<https://github.com/netwerk-digitaal-erfgoed/high-level-design/blob/master/building-blocks.md#service-portal>)

Ruben Verborgh: *“Decentralization needs replication for realistic performance”*

“...transparent layers in network of nodes...”

(<http://slides.verborgh.org/ELAG-2018/#>)

Central Platform: Current situation





http://data.adamlink.nl/

The screenshot shows a web browser window with the URL <https://data.adamlink.nl>. The page features a search bar at the top and four dataset cards arranged in a 2x2 grid. Each card includes a title, a description, and a grid of image thumbnails.

- Heritage** (by AdamNet, 8,744,254 statements): Description: "This dataset is a combination of all the heritage datasets available on data.adamlink.nl. It enables a user to find and use cultural heritage objects without the borders between institutions. The data are linked to streets, buildings, districts and persons they have in common as much as possible." Thumbnail: "Rembrandt's moeder aan tafel: driekwart lengte, naar rechts".
- Geography** (by AdamNet, 153,975 statements): Description: "This dataset contains all geographical concepts, collected by AdamNet, used to link the cultural heritage in the AdamLink dataset. It contains (old) streets, buildings and districts. See adamlink.nl." Thumbnail: "Rembrandthuis", "Jodenbreestraat", "Buurt 0".
- Images** (by Stadsarchief Amsterdam, 5,576,957 statements): Description: "This dataset contains the images from 'beeldbank' of the Amsterdam City Archives. Its linked to streets, buildings, districts and persons as much as possible." Thumbnail: "00-7008", "BMAA000010000", "B000000022466".
- Collection** (by Amsterdam Museum, 1,779,914 statements): Description: "This dataset is based on the catalogue of the Amsterdam Museum. It contains metadata on the objects in the collection, like painting, etches and utensils. For AdamLink they are linked to other cultural heritage objects in Amsterdam by means of AdamLink uri's for streets, buildings, districts and persons." Thumbnail: "Rembrandt's moeder aan tafel: driekwart", "Snijbetsel (graaftand) van 'Victoria'", and a painting thumbnail.

Local workflows

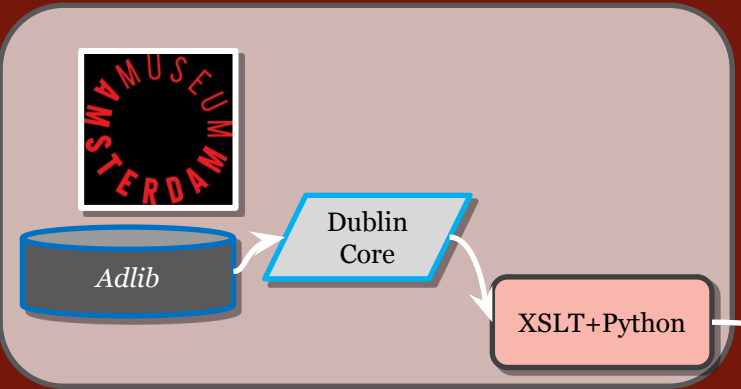
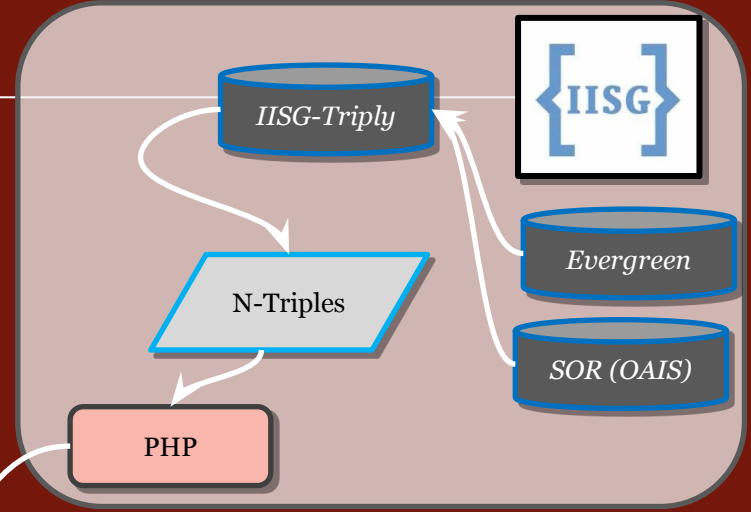


× Gemeente
× Amsterdam
× Stadsarchief

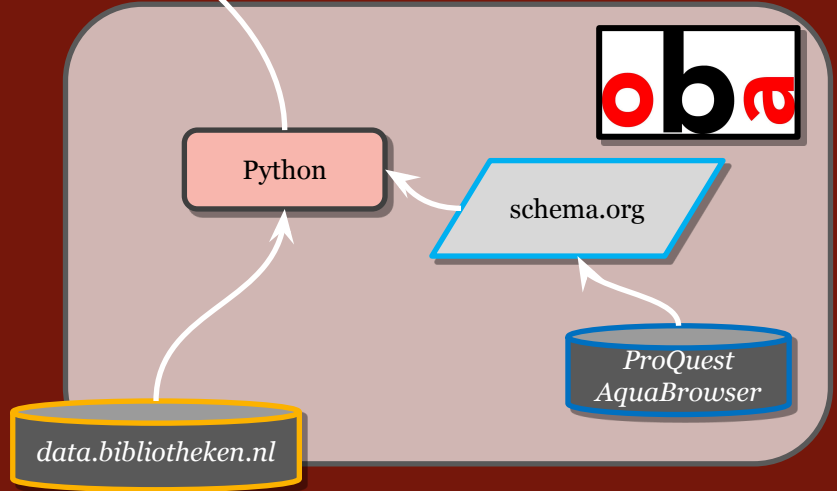
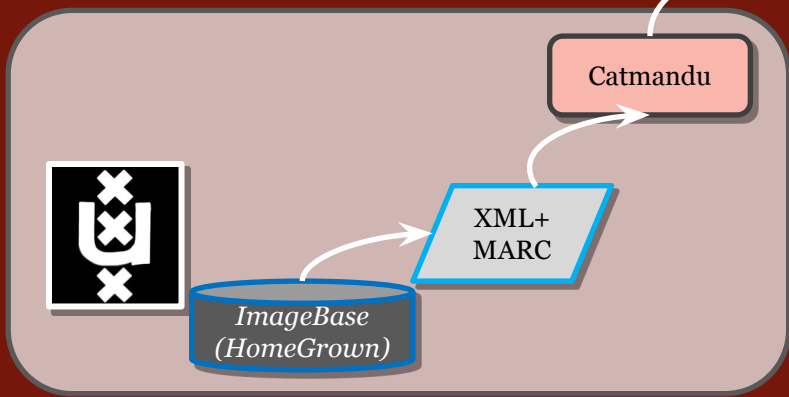
Memorix
(Picturae)

Dublin Core
+
Proprietary

PHP



Triply
RDF
Turtle





adamlink.nl: "linking points"

Locations

- Streets
- Buildings

Creators/contributors

- people
- organisations

Types

Subjects

- including
- people
- organisations

Jodenbreestraat edit

De Jodenbreestraat is een Amsterdamse straat tussen de Sint Antoniesbreestraat en het Mr. Visserplein in Amsterdam Centrum. De Jodenbreestraat is via de Houtkopers-dwarsstraat verbonden de met de noordzijde van het Waterlooplein. Ter hoogte van nr. 11 ligt de Uilenburgersteeg, meer op nl.wikipedia.org/wiki/Jodenbreestraat

Alle naamvarianten:

- Brede Straat
- Bredestraat
- Breestraat
- J. breestr.
- J. Breesstr.
- Jbreestr.
- Jod Breestraat
- Jod. Breestr.
- Jod. Breesstr.
- Jod.Breestraat
- Jodebreestraat
- Joden Bree straat
- Joden Breestr
- Joden Breestr.
- Joden Breestraat
- Jodenbeestr.
- Jodenbreestr
- Jodenbreestr.
- Jodenbreestraat
- Jodenbreeststraat
- Jodenbreetsr.
- Jofdenbreestr.

Identifiers

BAG id	0363300000003480
BAG URI	http://bag.basisregistraties.overheid.nl/bag/id/openbare-ruimte/0363300000003480
WikiData id	Q2689416
WikiData URI	http://www.wikidata.org/entity/Q2689416
Adamlink URI	https://adamlink.nl/geo/street/jodenbreestraat/2158

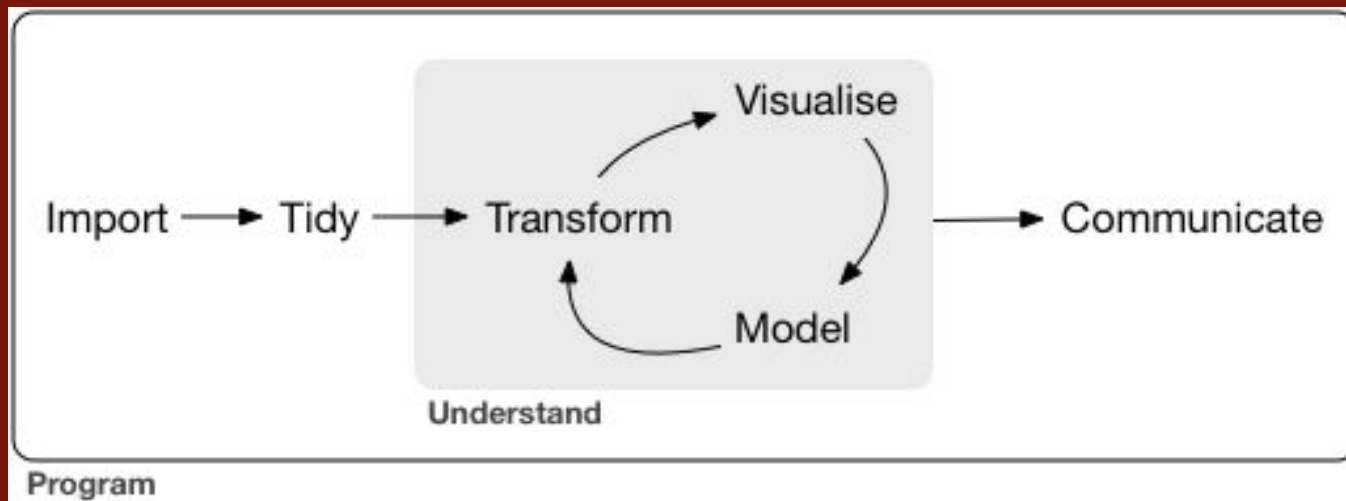
Geometrie

```
[{"type": "MultiLineString", "coordinates": [[[ [4.9024405140237, 52.369118760834], [4.9040422691128, 52.369400088948] ], [ [4.9009446555666, 52.369660132891], [4.9013281364349, 52.369800132451] ] ] ]}
```

Meer gegevens, waaronder de geometrie als MultiPolygon (in RD) te vinden op <https://www.openstreetmap.org/way/111111111>

What we need is *transformation*

“Aggregation” into one dataset at one endpoint



- <https://r4ds.had.co.nz/introduction.html>

Transformations

1. Ontology alignment
2. Authority alignment
3. Object types alignment
4. Additional statements
5. Restructuring data
6. Data typing

Example: Herengracht / Springer, C.





Example: Herengracht / Springer, C.

```
<http://hdl.handle.net/11259/collection.37414>
  rdf:type          edm:ProvidedCHO ;
  rdfs:label        "De bocht in de Herengracht"^^xsd:string ;
  dc:identificier   "SA 286"^^xsd:string ;
  dc:title          "De bocht in de Herengracht"^^xsd:string ;
  sem:hasBeginTimeStamp "1882"^^xsd:string ;
  sem:hasEndTimeStamp  "1882"^^xsd:string ;
  foaf:depiction    <http://amasp.adlibhosting.com/wwwopacx_images...>;
  dc:creator        <https://adamlink.nl/person/springer-cornelis/2220> ;
  dc:subject        "wintergezicht"^^xsd:string ;
  dc:type           <http://vocab.getty.edu/aat/300177435> ;
  dct:spatial       <https://adamlink.nl/geo/street/herengracht/1768> ;
  edm:isShownAt    <http://hdl.handle.net/11259/collection.37414> ;
  dc:rights        <http://creativecommons.org/publicdomain/mark/1.0/> ;
  void:inDataset   <https://data.adamlink.nl/am/collection/> .
```



(1/6) Ontology alignment: Common Data model?

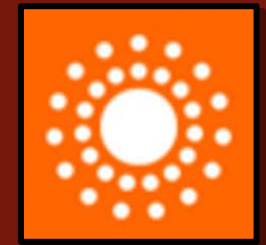
≡ Europeana Data Model?

- ≡ Targeted at data aggregation/consolidation
- ≡ Focused on internal Europeana workflow
- ≡ Mainly DC properties
- ≡ EDM properties: internal procedures



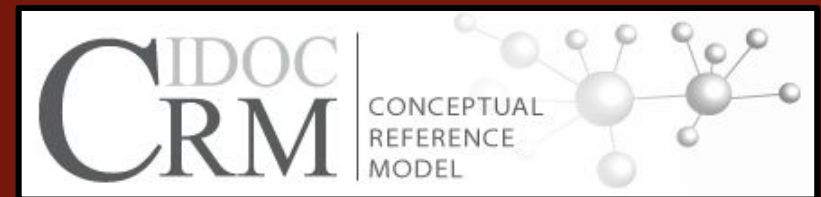
≡ Dublin Core?

- ≡ Not rich enough











≡ CIDOC-CRM?

- ≡ Not simple enough



(1/6) Ontology Alignment: Common properties

- ☰ rdf:type 
- ☰ foaf:depiction 
- ☰ rdfs:label 
- ☰ dc:creator  schema:author, ...
- ☰ dc:contributor 
- ☰ dc:type 
- ☰ dc:subject  schema:about, ...
- ☰ dct:spatial 
- ☰ dct:temporal
- ☰ dc:date
- ☰ sem:hasBeginTimeStamp
- ☰ sem:hasEndTimeStamp
- ☰

(2/6) Authority alignment: Common vocabularies



Type

- ☰ AAT



People

- ☰ VIAF
- ☰ WikiData
- ☰ RKD-Artists
- ☰ ...



Subjects

- ☰ WikiData
- ☰ (spatial) Geonames, TGN
- ☰ ...





(2/6) Authority alignment: sameAs

- ≡ types - AAT URI
- ≡ creator/contributor -> AdamLink URI
- ≡ subject (for persons) -> AdamLink URI
- ≡ location -> AdamLink URI



(3/6) Object types alignment

rdf:type:

- ≡ edm:ProvidedCHO
- ≡ schema:Person, schema:Organization
- ≡ hg:Street, hg:Building, hg:District



(4/6) Additional statements

Data management metadata

organisations, datasets

- ≡ void:inDataset - AdamLink dataset URI
- ≡ dcterms:publisher - AdamLink Org URI
- ≡ dc:type - AAT URI, derived from rdf:type



(5/6) Restructuring data

eg. From EDM to DC:

- ≡ Rijksmuseum data
- ≡ complex path (via `ore:Aggregation-edm:WebResource`) between `object-uri` and the location of an image
- ≡ transformed into `foaf:depiction`



(6/6) Data typing

Complex, not fully developed:

xsd:string, xsd:int, xsd:date

or should this be the responsibility of data provider?



Conclusion

- ≡ decentral vs central
- ≡ aggregation or not
- ≡ transformation

... all depend on the use case!

We tried to develop an overview of the steps needed in transformation

Discussion: who is responsible for what?



That's all Folks!

<https://tinyurl.com/adamlinkSWIB18>