

Publishing linked data on data.bibliotheken.nl

René Voorburg

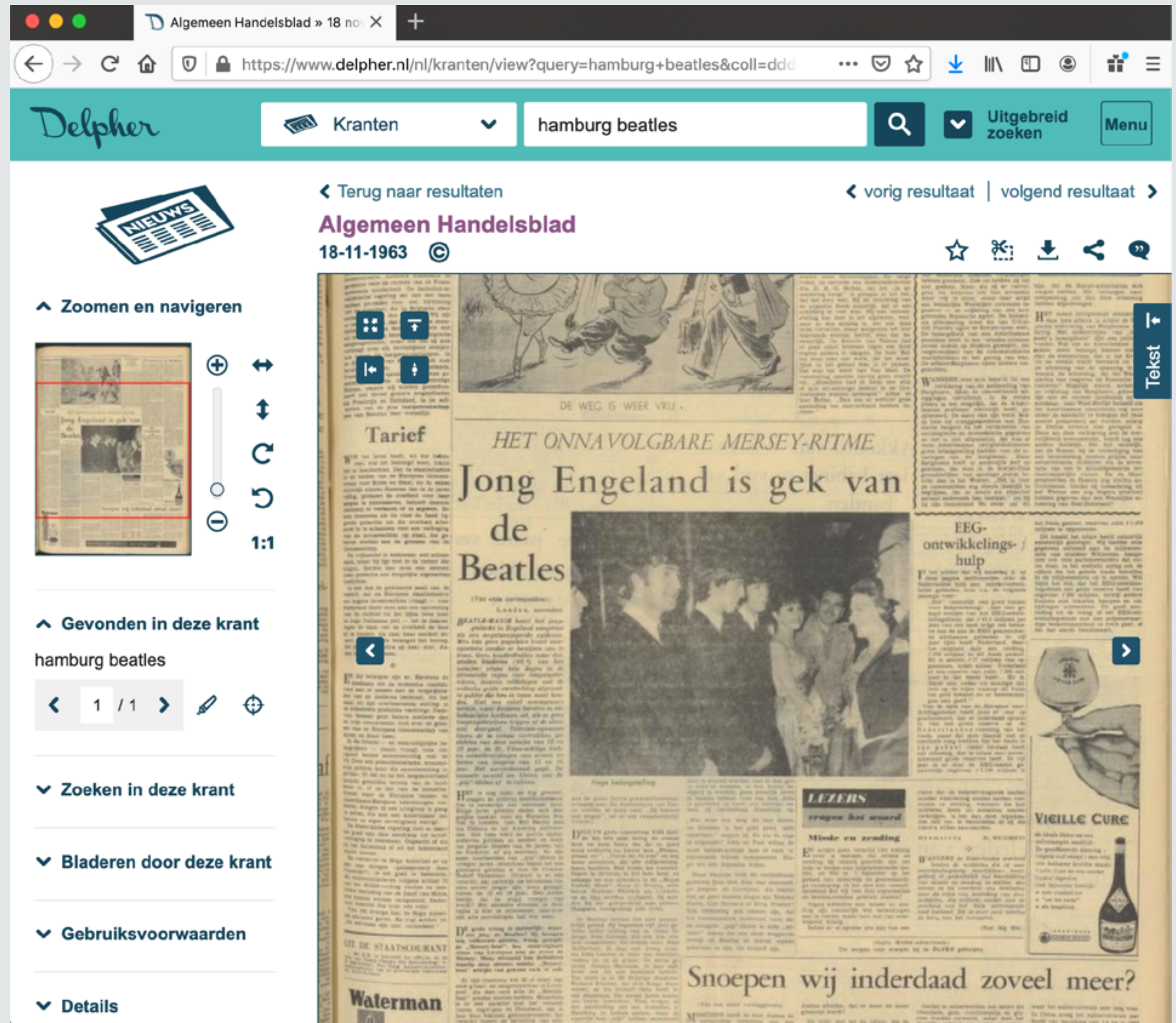
rene.voorburg@kb.nl

KB metadata infrastructure

- > 4.5 mln catalogue records
- ~ 1.5 mln digitized newspapers
- ~ 135 mln newspaper articles
- ~ 300.000 digitized periodicals
- ~ 900.000 digitized books

‘Organically grown’ metadata

Provided through OAI-PMH, SRU (SOLR), http-proxy (resolver).



Why start publishing linked data?

“To be in the web, instead of just on the web”

Facilitating others to use, reuse and link to KB metadata (and digital objects).

Core semantic & design principles

- Start with bibliographic metadata, the foundation of all
- IFLA LRM ontology (Work - Expression - Manifestation - Item) defines the horizon
- Map to schema.org as much as possible / when sensible
- De-conceptualize (/de-SKOS) whenever beneficial
- URI policy: distinguish between the thing and its description ('303 see other')

Technical foundations

- LodView RDF viewer
- Virtuoso Open Source triple store backend for LodView provides SPARQL interface
- Apache webserver for configuring redirects, content negotiation, etc

Some of the approaches followed in creating the RDF

Thesauri:

Pica+

XSLT

skos-RDF

SPARQL

schema-RDF

Set 'DBNL':

relational db

PHP

schema-RDF

Bibliographic records:

Pica+

XSLT

'Pica+'-RDF

SPARQL

schema-RDF

Improved approach:

Pica+

XSLT

relational db

stored procedure
XSLT

schema-RDF

http://data.bibliotheken.nl/

De toegang tot alle Linked Data zoals beschikbaar gesteld door de [Koninklijke Bibliotheek](#).
Deze dienst is een bètaversie. Meer informatie over deze en andere datasets en -services? Kijk bij [Dataproducten en API's](#).

Dataset	URI
"Alba amicorum van de Koninklijke Bibliotheek"@nl	http://data.bibliotheken.nl/id/dataset/rise-alba
"Brinkman trefwoordenthesaurus"@nl	http://data.bibliotheken.nl/id/dataset/brinkman
"Centsprenten, een goedkope vorm van massadrukwerk uit de achttiende en negentiende eeuw."@nl	http://data.bibliotheken.nl/id/dataset/rise-centsprenten
"Gemeenschappelijke Trefwoordenthesaurus (GTT)"@nl	http://data.bibliotheken.nl/id/dataset/gtt
"Nederlandse Bibliografie Totaal (NBT)"@nl	http://data.bibliotheken.nl/id/dataset/nbt
"Organisaties uit de corporatiethesaurus van de Koninklijke Bibliotheek"@nl	http://data.bibliotheken.nl/id/dataset/corps
"Personen uit de Nederlandse Thesaurus van Auteursnamen (NTA)"@nl	http://data.bibliotheken.nl/id/dataset/persons
"Short-Title Catalogue Netherlands (STCN)"@nl	http://data.bibliotheken.nl/id/dataset/stcn
"Thesaurus Auteurs DBNL"@nl	http://data.bibliotheken.nl/id/dataset/dbnla
"Thesaurus KBcode"@nl	http://data.bibliotheken.nl/id/dataset/kbcode
"Titels DBNL"@nl	http://data.bibliotheken.nl/id/dataset/dbnlt

Nederlandse Bibliografie Totaal (NBT)

http://data.bibliotheken.nl/id/dataset/nbt

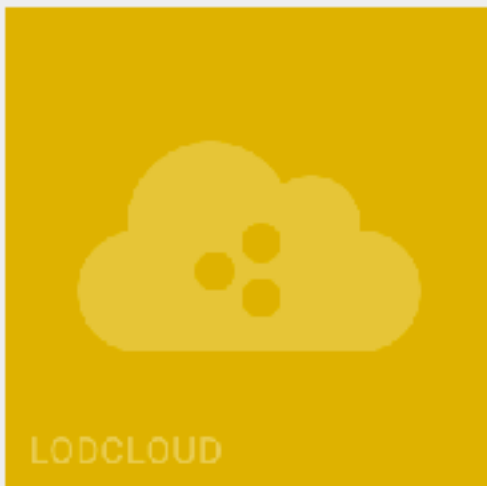
TYPE ENTITEIT: Dataset

rdfs:label	Nederlandse Bibliografie Totaal (NBT) @nl
cc:attributionName	Nederlandse Bibliografie Totaal (NBT) @nl
schema:description	Nederlandse Bibliografie Totaal (NBT) bevat publicaties die in Nederland, over Nederland of in de Nederlandse taal zijn verschenen @nl
schema:name	Nederlandse Bibliografie Totaal (NBT) @nl
rdf:type	void:Dataset
	schema:Dataset
rdfs:seeAlso	<https://www.kb.nl/bronnen-zoekwijzers/dataservices-en-apis>
	<https://www.kb.nl/organisatie/voor-uitgevers/informatie-over-de-nederlandse-bibliografie>
cc:attributionURL	<http://data.bibliotheken.nl>
cc:license	<http://opendatacommons.org/licenses/by/1.0/>
cc:morePermissions	<mailto:dataservices@kb.nl>
schema:publisher	<http://data.bibliotheken.nl/id/thes/p075301482> ↳ Koninklijke Bibliotheek (Den Haag)
void:exampleResource	<http://data.bibliotheken.nl/doc/nbt/p357539168>
void:sparqlEndpoint	<http://data.bibliotheken.nl/sparql>

Het stenen bruidsbed : roman / Harry Mulisch

<http://data.bibliotheken.nl/id/nbt/p357539168>

TYPE ENTITEIT: **Book**



rdfs:label	Het stenen bruidsbed : roman / Harry Mulisch
schema:isbn	9789023476825
schema:description	1e dr.: 1959. - (Literaire reuzenpocket ; 1)
schema:bookEdition	44e dr
schema:inLanguage	ned
schema:name	Het stenen bruidsbed : roman
schema:numberOfPages	183 p
rdf:type	schema:Book
owl:sameAs	< http://www.worldcat.org/oclc/844922907 >
schema:author	< http://data.bibliotheken.nl/id/thes/p06854796X > ↳ Mulisch, Harry (1927-2010)

Some things we 've learned and want to improve

Creating and updating RDF is a lot of work

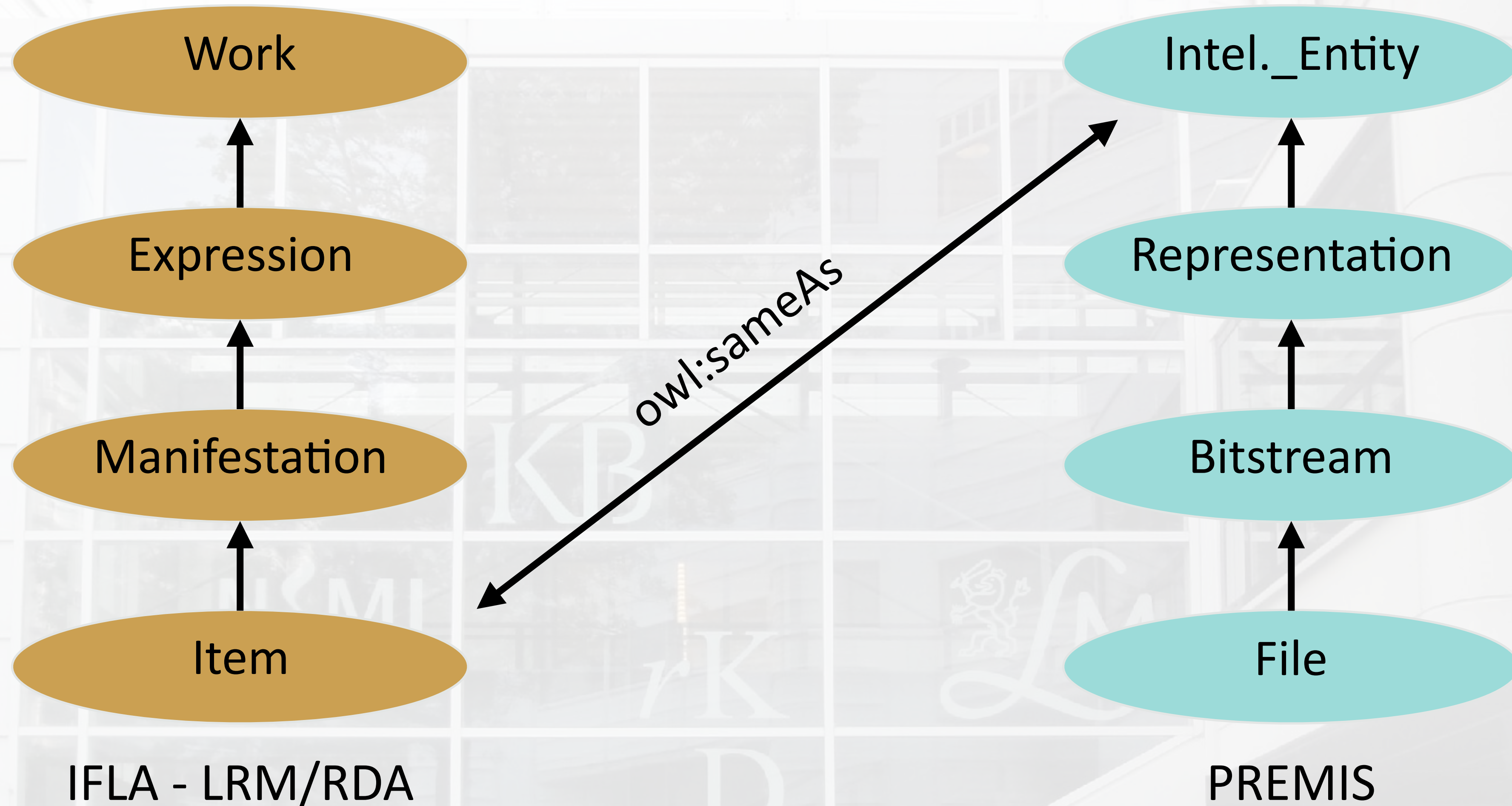
- More automation is wanted
- Ideally, core metadata systems should be linked data compatible

Ideally, all KB metadata should adhere to one generic entity relationship model

- to answer the core question “exactly what are we describing here”

For different uses, ‘views’ with different levels of detail are wanted

Towards a unified KB metadata model



LRM/RDA entity clustering

Experimental LRM /RDA clustering
at LRM:Work and LRM:Expression
level using the 'RDA entity finder'.

Part of ongoing work to map Pica+
to RDA.

KB } nationale
bibliotheek

RDA Entity Finder

Titel

de aanslag

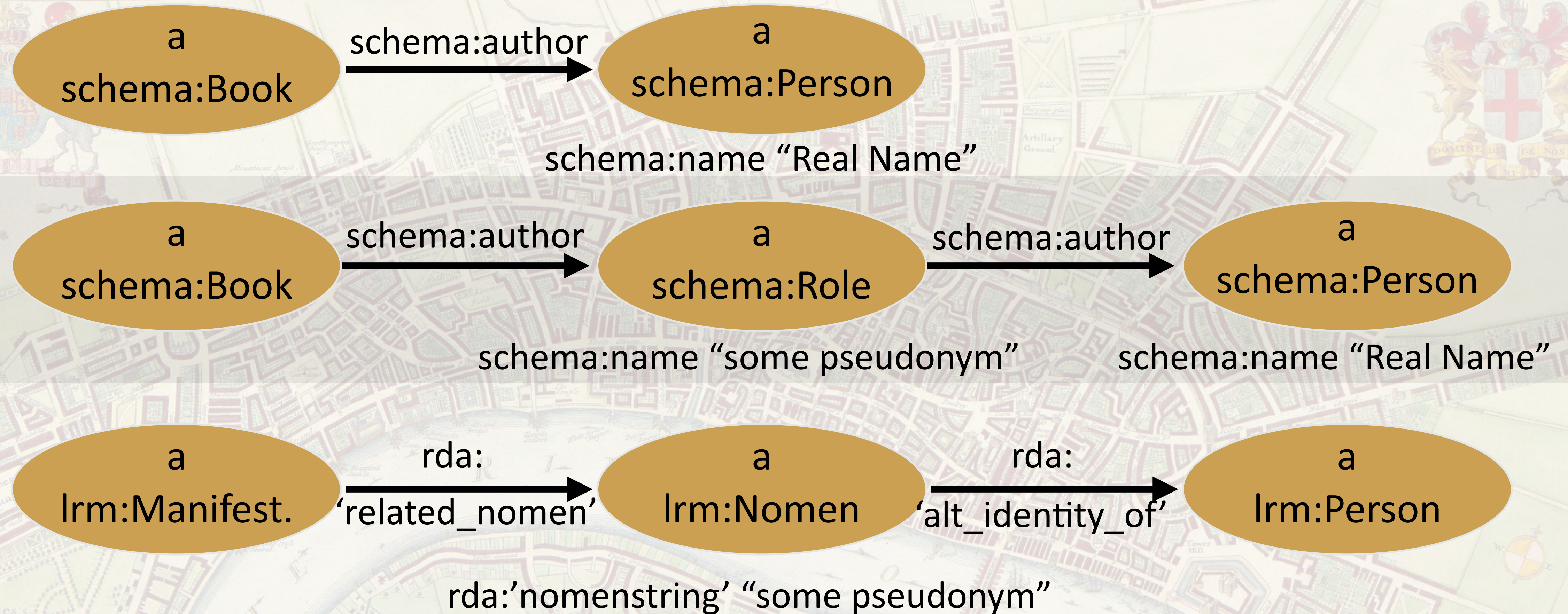
Auteur

mulisch

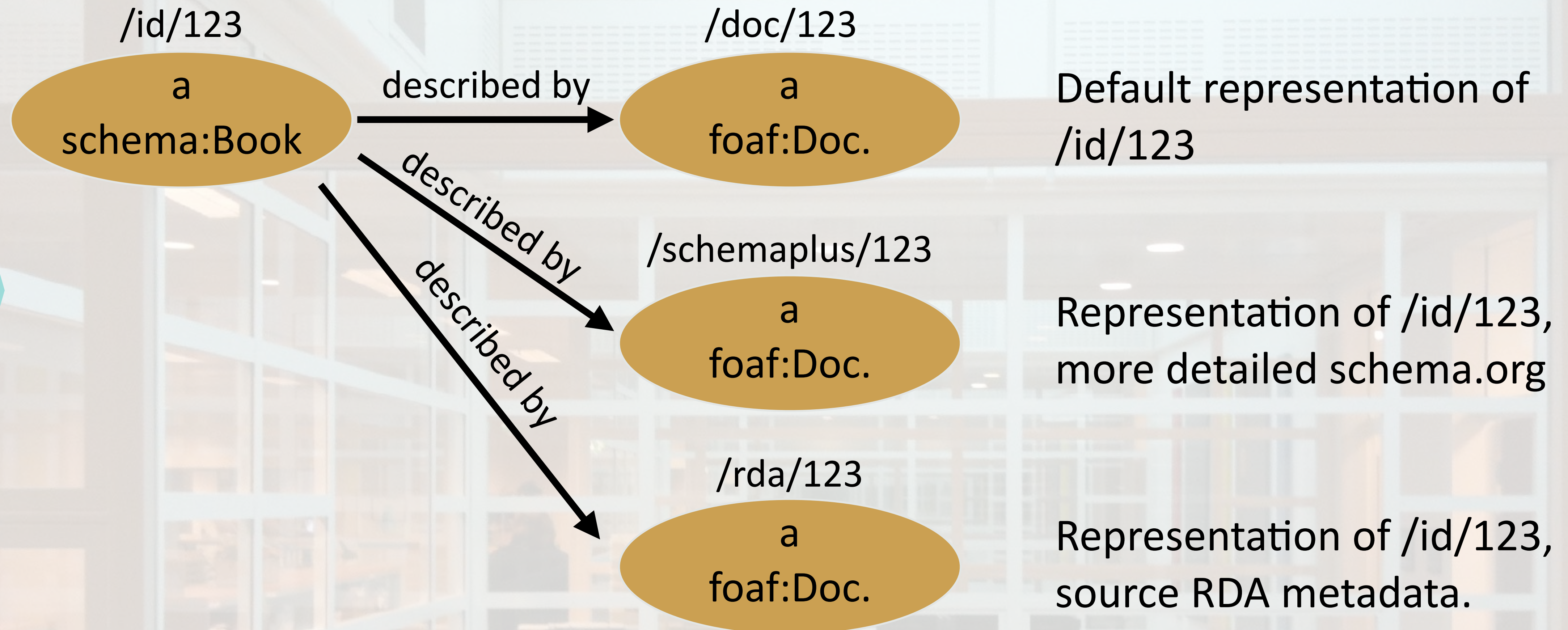
Zoeken

Work #1: de aanslag - Mulisch, Harry Kurt Victor ; Mulisch, Harry Kurt Victor	Details tonen	Expressies inklappen
Expression #1: gesproken woord, ned	Details tonen	Manifestaties tonen
Expression #2: tekst, bul, Atentatāt	Details tonen	Manifestaties tonen
Expression #3: tekst, dee, Et tilfælde	Details tonen	Manifestaties tonen
Expression #4: tekst, dui, Das Attentat	Details inklappen	Manifestaties tonen
has content type: text has language of expression: German has contributor: Habers, Annelen is translation of: \W1\mulisch\W2\de aanslag\E1\rdaco:1020\E2\ned\ title of expression: Das Attentat authorized access point for expression: \W1\mulisch\W2\de aanslag\E1\rdaco:1020\E2\dui\E4\das attentat\ PPN: 861720962		
Expression #5: tekst, eng, White, Claire Nicolas, The assault	Details tonen	Manifestaties tonen
Expression #6: tekst, fra. Noble. Philippe. L'attentat	Details tonen	Manifestaties tonen

Data on the move, a case for multiple representations



Modelling multiple representations



To conclude

- The linked data effort that started off with an outward focus ('the user'), has now turned inwards ('how to improve our metadata and metadata infrastructure').
- Linked data has worked as a catalyst for cooperation between the traditional metadata department (the catalogue) and the more IT-related metadata-services departments.
- With regards to tools, standards, vocabularies & ontologies, there is still much work to be done to reach maturity.