



Data op het Web

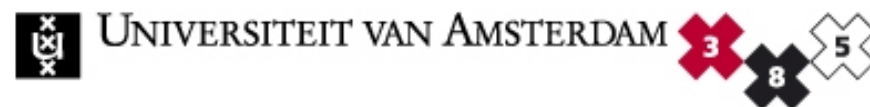
Amsterdamse
instellingen verbeteren
toegang tot hun collecties
over de Amsterdamse
geschiedenis ...

... met behulp van Linked
Open Data.

Dankzij:



× Gemeente
× Amsterdam
× Stadsarchief





Inhoud

- Hoe verbind ik mijn data aan het Web van Data? (**TED-talk TBL**)
- Voor wie: collectiebeheerders





Actie!

- 1) Verbeter de data
- 2) Publiceer de data
- 3) Bied data-services

Half
is
beter
dan
niet!





(1) Verbeter de data

- a) Leg rechten vast
- b) Introduceer URI's
- c) Link de data



(1a) Leg rechten vast

- Elk object heeft zijn eigen rechten
- Kies een standaard en leg de rechten vast met een verwijzing naar het auteursrecht
- Probeer zoveel mogelijk materiaal in het Publieke Domein te krijgen





1b) Introduceer URI's

- Elk object krijgt zijn eigen webadres
- Dit webadres blijft voor altijd werken



– bv. <http://hdl.handle.net/11259/collection.277>



(1c) Link de data

- Gebruik een standaard voor het beschrijven
 - neem bv. het RKD nummer “66219” op in de beschrijving in plaats van de string “R(h)ijn, R.H. v.” of zoiets.
- Gebruik nog liever uri's
 - <https://rkd.nl/explore/artists/66219>



(1c) Link de data

- Dit kan met:
 - straten, adressen en bijzondere gebouwen
 - personen (vervaardigers, maar ook geportretteerden)
 - documenttypes
 - auteursrechten
 - andere erfgoedobjecten
 - etc. etc.





Tussenstand

- De data is veel beter: gestandaardiseerde waarden, unieke identifiers, rechten uitgezocht en vastgelegd.
- ... en we hebben nog niets online staan :-)
- Je mag het geleidelijk doen: eerst een paar, dan uitbreiden.





(2) Publiceer de data

- a) Zet de dataset online
- b) Zet de dataset online in RDF-formaat
- c) Publiceer de data per object



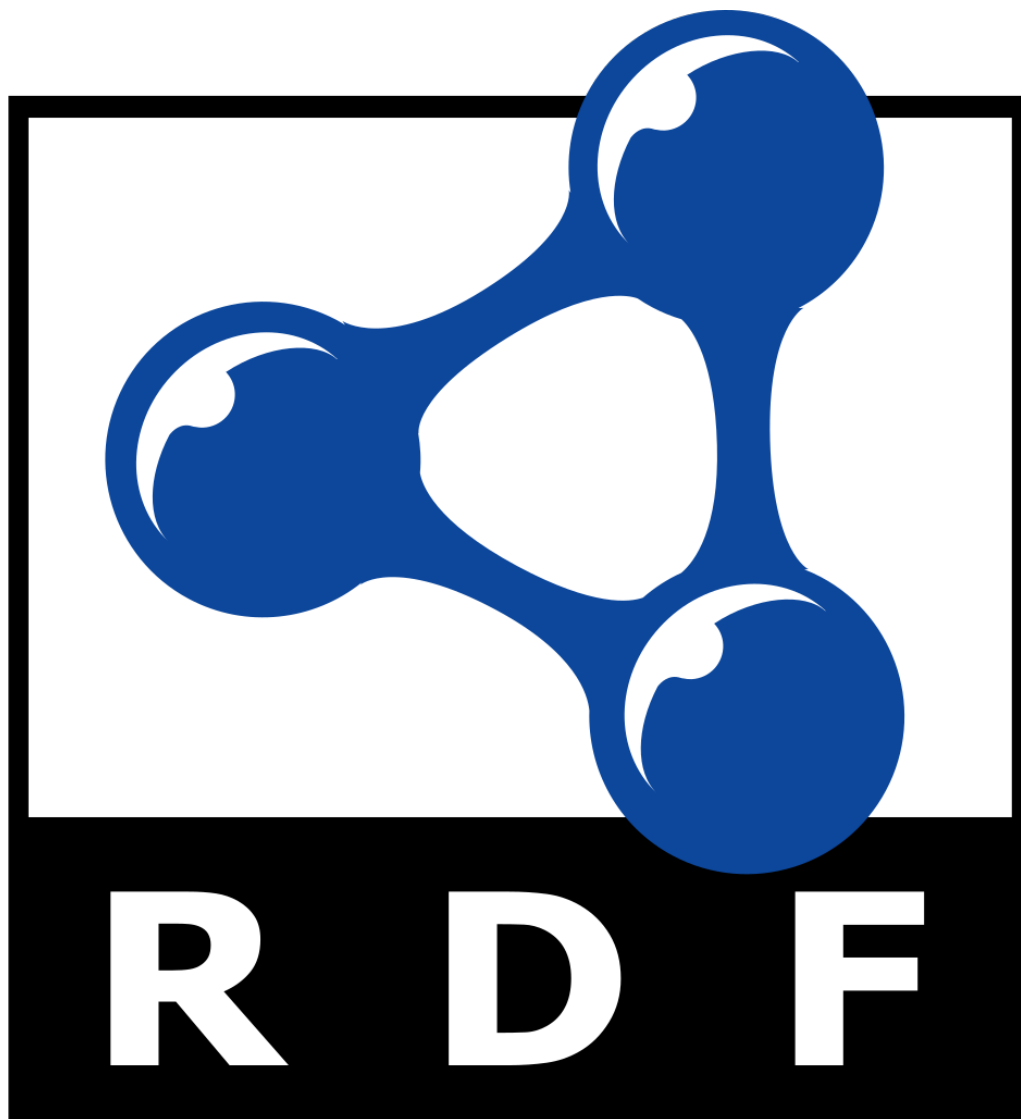
(2a) Dataset online

- Maak een webpagina met een link naar het data-bestand
- Als excel-sheet, in CSV, TXT, XML formaat ...
- ... maakt niet uit: alles is beter dan niets.
- Bij voorkeur geautomatiseerd



(2b) RDF-formaat

Resource
Description
Framework





(2b) Triples

<ets>

<is gemaakt door>

<Rembrandt>



<ets> = <<http://hdl.handle.net/11259/collection.277>>

<is gemaakt door> = dc:creator

<Rembrandt> = <<https://rkd.nl/explore/artists/66219>>



(2b) Serialisaties

Er zijn verschillende
verschijningsvormen
van RDF: RDF/XML, nq.

Turtle:

<<http://hdl.handle.net/11259/collection.277>>

dc:title "Eene zittende oude vrouw (Rembrandt's moeder)" ;

dc:creator <<https://rkd.nl/explore/artists/66219>> ;

dc:rights <<http://creativecommons.org/publicdomain/mark/1.0/>> ;

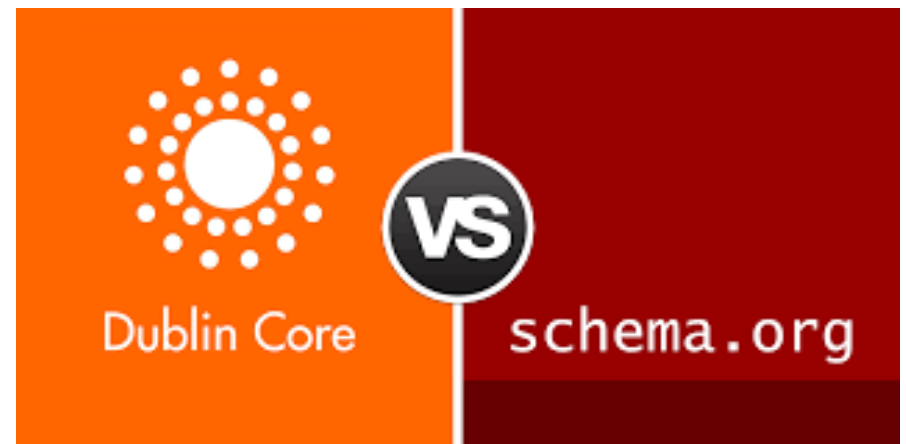
dc:subject <<http://www.vondel.humanities.uva.nl/ecartico/persons/8953>> .





(2b) RDF-formaat

- Dus:
 - Kies bestaande properties
 - (bijvoorbeeld uit Dublin Core)
 - Programmeer een export naar RDF
 - (bijvoorbeeld in turtle)
 - Zet een link naar het bestand in RDF formaat op je data-webpagina

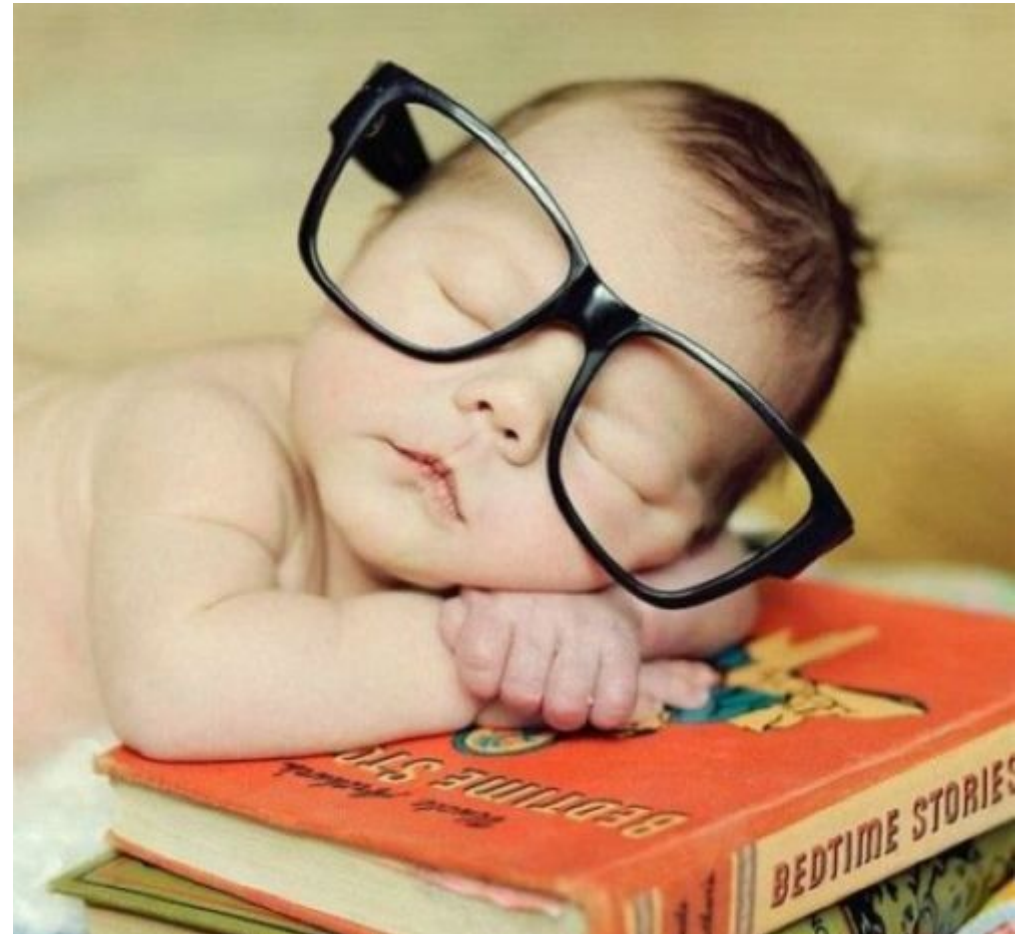




(2c) Data per object

Als je data wil ophalen
over een specifiek
object:

- Resolvable uri's
- Content Negotiation





Tussenstand

Waarom is RDF beter dan Excel, CSV, TXT, XML etc.?



(3) Bied data-services

- a) Synchronisatie mechanisme
- b) Application Programming Interface
- c) SPARQL-endpoint



(3a) Synchronisatie

Gebruikers willen misschien niet elke keer alle data helemaal opnieuw downloaden en verwerken, maar alleen de gewijzigde data.

- Daarvoor kun je een synchronisatie-dienst aanbieden:
 - OAI-PMH: OAI Protocol for Metadata Harvesting
 - Resource Sync: sitemaps



(3b) API

Webontwikkelaars zijn gewend aan een Application Programming Interface (API). Daarmee schrijf jij voor welke mogelijkheden er zijn om met de data om te gaan.

- Resultaten worden meestal in “JSON” geleverd.



(3c) SPARQL-endpoint

Een SPARQL endpoint geeft de volledige rijkdom aan mogelijkheden met de data. Een gebruiker kan dan een zoekvraag stellen, heel vergelijkbaar met de database query-taal SQL.

Score-card

Verbeter de data	Leg rechten vast	Introduceer URI's	Link de data
Publiceer de data	Zet dataset online	Zet dataset online in RDF-formaat	Publiceer data per object
Bied data-services	Synchronisatie mechanisme	Application Programming Interface	SPARQL-endpoint

Linked Open Data (★★★★★)

Verbeter de data	 Leg rechten vast	 Introduceer URI's	 Link de data
Publiceer de data	Zet dataset online	 Zet dataset online in RDF-formaat	 Publiceer data per object
Bied data-services	Synchronisatie mechanisme	Application Programming Interface	SPARQL-endpoint



Conclusie

- Belangrijkste is het verbeteren van de data.
- Om te beginnen heb je geen techneut nodig: je vervangt de waardes in losse invulvelden door uri's.
- Je kunt klein beginnen en dan uitbouwen.