

HANDS-ON MET HET DELVING LINKED OPEN DATA PUBLICATIE PLATFORM

~ sjoerd Siebinga (Delving) ~

WIE ZIJN WIJ

- Delving BV is in 2010 ontstaan uit het core team van Europeana
- Vooral overheid, semi-overheid en cultuur-instellingen als klanten
- Het hele platform is Open Source
- core-business is metadata aggregatie en RDF transformatie

AGENDA

- Linked Open Data life-cycle
- Ingestion Workflow Hub3
- Verrijking met de Erfgoed Thesaurus

LOD IS TOCH NIET MOEILIJK

- pak wat data
- pak wat vocabulaires
- maak wat RDF
- voeg wat LOD links toe
- stop het in een Triple Store

EN ... KLAAR !

... MAAR DAN ...

... DE REALITEIT!

ONZE ACHTERGROND

- 4 jaar productie
- 1000+ datasets
- miljarden triples
- de meeste worden dagelijks geüpdate
- brondata worden geaggregeerd en getransformeerd

BRON-DATA VS AGGREGATIE

bron-data publicatie

heeft

andere uitdagingen dan aggregatie

WELKE VALKUILEN ZIJN ER DAN?

versies en verweesde data

keuze van tools en hosting

welke niveau van kennis is voldoende

hoe verrijkingen te maken en op te slaan

automatische publicatie pijplijn

hoe beschrijf je je RDF dataset

cachen van externe data en schemas

hoe hou je alles up-to date

validatie van je URIs en gepubliceerde RDF

...

WAT MOET JE DOEN?

- kies een goede en duurzame base-url
- hoe ga je data publiceren
 - LOD content negotiation
 - SPARQL
 - Linked Data Fragments
 - HDT
- hoe ga je beschikbaarheid garanderen

WAT HEB JE NODIG?

- analyse van je eigen data
- transformatie/mapping tool naar je target formaat
- duurzame opslaglaag voor RDF met versie beheer
- tools om je data te verrijken met externe bronnen
- compliance met W3C standaarden, zoals:
 - SPARQL
 - RDF

WAT BIEDT DE HUB₃ IN EEN NOTENDOP

- graphische mapping ingestion workflow
- RDF upload endpoints
- full-tekst en geo-zoek API
- beheer van datasets in DCAT
- beveiligde SPARQL proxy
- IIF compliant media-manager met DeepZoom
- verrijkingen dmv SKOS thesauri
- automatisch cachten en up to date houden van externe bronnen
- ... en nog veel meer

DEMO TIME

- data ingestion
- mappen naar EDM
- online publiceren
- review endpoints
- verrijken van literals met SKOS