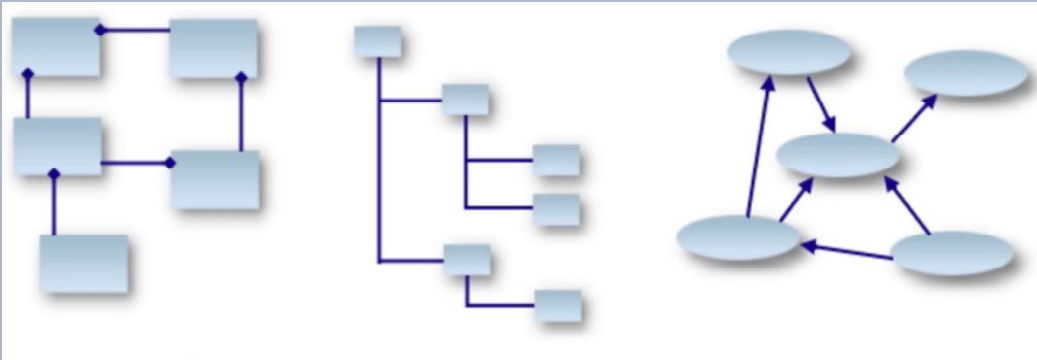




Linked Data en GEO Ruimteschepper

Martien Vos
Woensdag 3 Juli 2024
martenvos@gmail.com

Database modellen



Relationale DB

- Relaties via Primäre sleutels

Hierarchische DB

- Gebaseerd op Master nodes

Graph DB

- Willekeurige objekt relaties

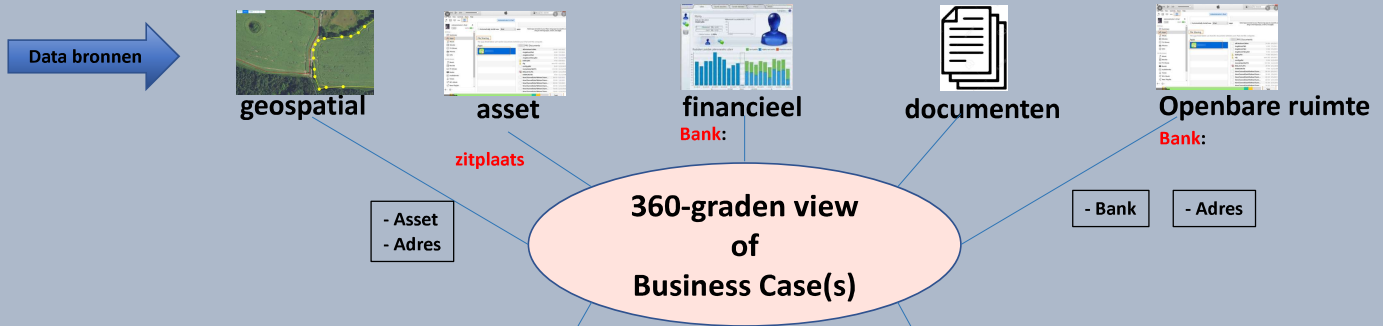
Soorten Graphs

- Verschillende toepassingsscenario's
 - **RDF Semantic Graphs** (Resource Description Framework)
gegevens modelleren als triples, semantische toepassingen (betekenis toevoegen)
 - **Property Graphs**
K/V-paren (attributen) associeerd met vertexes/edges voor analyse van b.v. sociale netwerken, financiële data
 - **Link-node graphs**
fysieke/logische netwerken vertegenwoordigen die worden gebruikt in o.a. transport, nutsbedrijven en telecom (traditionele GIS - Topologie)

Waarom kiezen voor een RDF Graph?

- Realiseert data-integratie zonder continue fysieke datarePLICatie
- Maakt gebruik van internet technologie
 - Betaalbaar & alomtegenwoordig (overal verkrijgbaar)
 - Iedereen is verbonden via internet
 - Dominante technologie voor het verbinden van gegevens
 - Kan verbinding maken met elke gegevensbron mits toegekende autorisatie
- Real-time gegevens ophalen uit de bestaande gegevensbronnen
- Maakt smart geospatial objecten mogelijk.
- Gebaseerd op de open standaarden van W3C en OGC
 - Kennismodellen kunnen worden geïntegreerd of verbonden met andere partijen
 - Maakt integratie van geospatial gegevens mogelijk (GeoSPARQL)

Data integratie problemen oplossen



Vergelijken op attributen

- Bank: →
- bank / zitplaats →
-

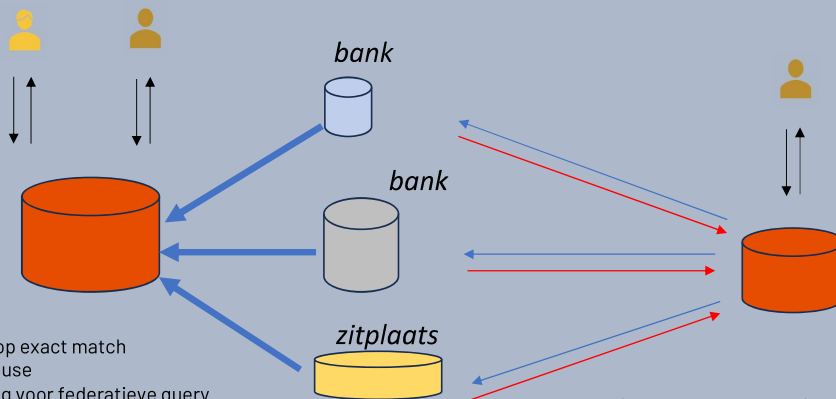
Probleem	Linked Data
Zelfde naam, verschillende betekenis	betekenis (context)
Verschillende namen, zelfde betekenis	betekenis (context)
Verschillende structuren	structuur
Resultaat	
Hoge kosten	Lagere kosten
Veel onderhoud	Minder onderhoud
Complex nieuwe business case	Sneller nieuwe toepassingen
Niet flexibel	Zeer flexibel

Vergelijken op betekenis

- bank = financiële instelling
- Bank = parkbank
- Bank - (parkbank) = zitplaats
- Bank → - Adres
- Asset →

360° inzicht halen uit meerdere gegevensbronnen

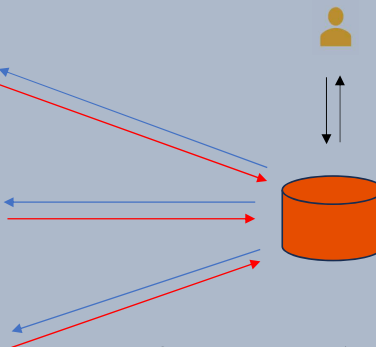
Vaak toegepaste methode



- Gebaseerd op exact match
- Datawarehouse
- Minder nuttig voor federatieve query vanwege mismatch en verbindingsproblemen
- Uiterst kwetsbaar voor de mogelijke drift van datamodellen aan de bronzijde

Linked Data methode.

Virtuele laag van betekenis (semantiek) over bestaande structuren



- Gebaseerd op semantisch matching (betekenis)
- Maakt gebruik van internet technologie
- FAIR data (findability, accessibility, interoperability, and reusability)
- Geschikt voor federatieve query
- Ook geschikt voor Geo data

Linked Data met Geo

- De basis eenheid van informatie is een Triple (Onderwerp – Gezegde – Lijdend voorwerp)
- Hiermee kunnen uitspraken worden gedaan over zaken in de werkelijkheid Dus over mensen, plaatsen, concepten enz. Maar ook over abstracte concepten en thema's.
- Tezamen vormen de triples een "Graph"

Bewering: Wageningen heeft 42.616 inwoners

Bewering: Wageningen behoort tot provincie Gelderland

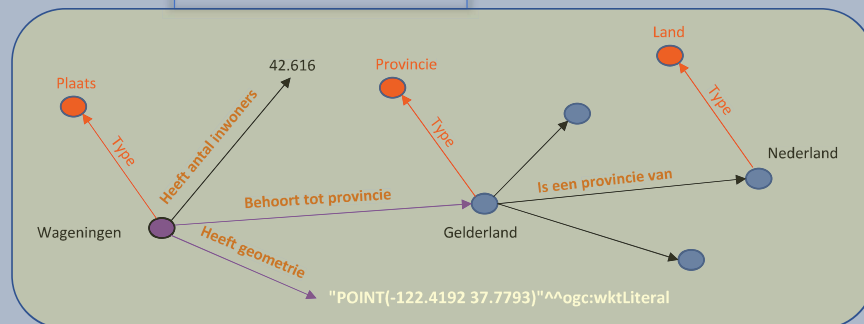
Bewering: Gelderland is een provincie van Nederland

Bewering: Wageningen is van type plaats

Bewering: Wageningen heeft als geometrie Point

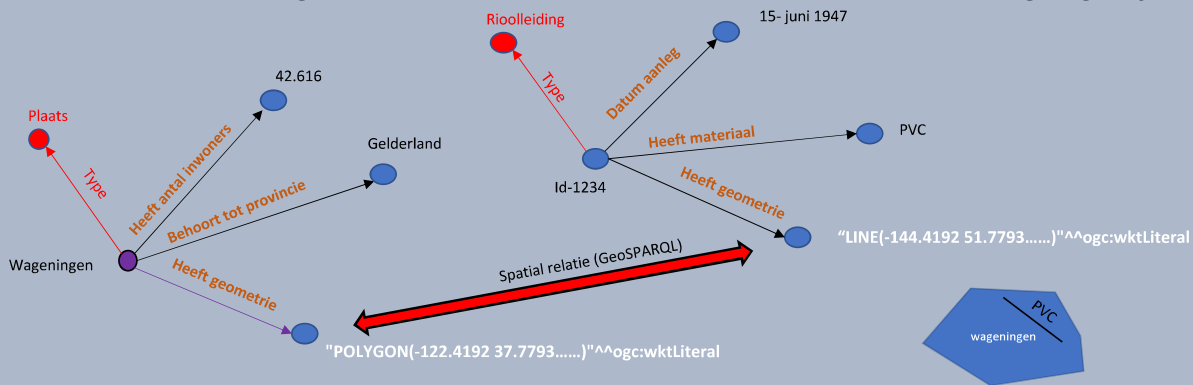


Voorbeeld RDF Graph

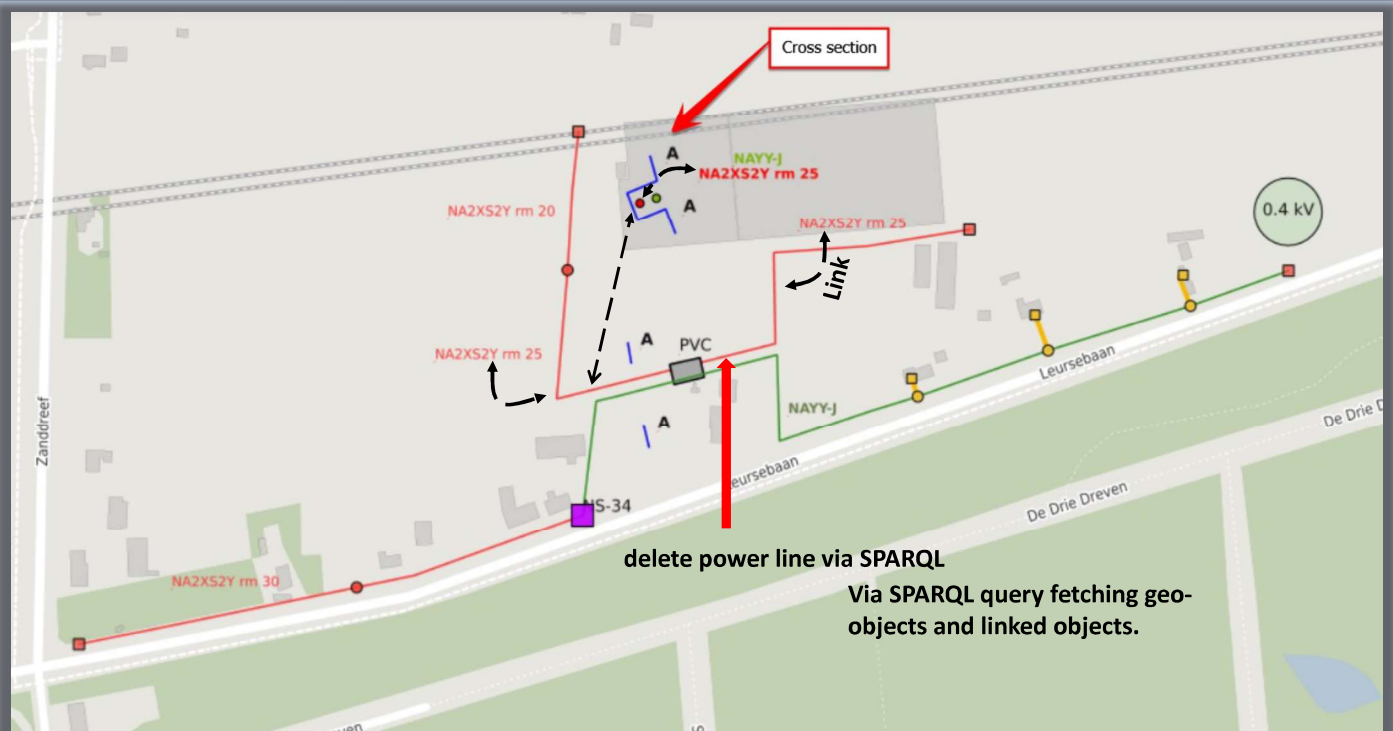


Bevragen van Linked Data met Geo (RDF)

- SPARQL = **SPARQL Protocol and RDF Query Language**. (W3C)
- Zoek alle plaatsen in Gelderland die meer als 40.000 inwoners hebben.
- Zoek rioolleidingen die als materiaal PVC hebben en die aangelegd zijn voor 1948
- GeoSPARQL = **extentie speciaal voor geo data**. (OGC)
- Zoek alle plaatsen in Gelderland die meer als 40.000 inwoners hebben en waar zich rioolleidingen bevinden die als materiaal PVC hebben en die aangelegd zijn voor 1948



Linked Data en Smart – GEO (data kwaliteit)

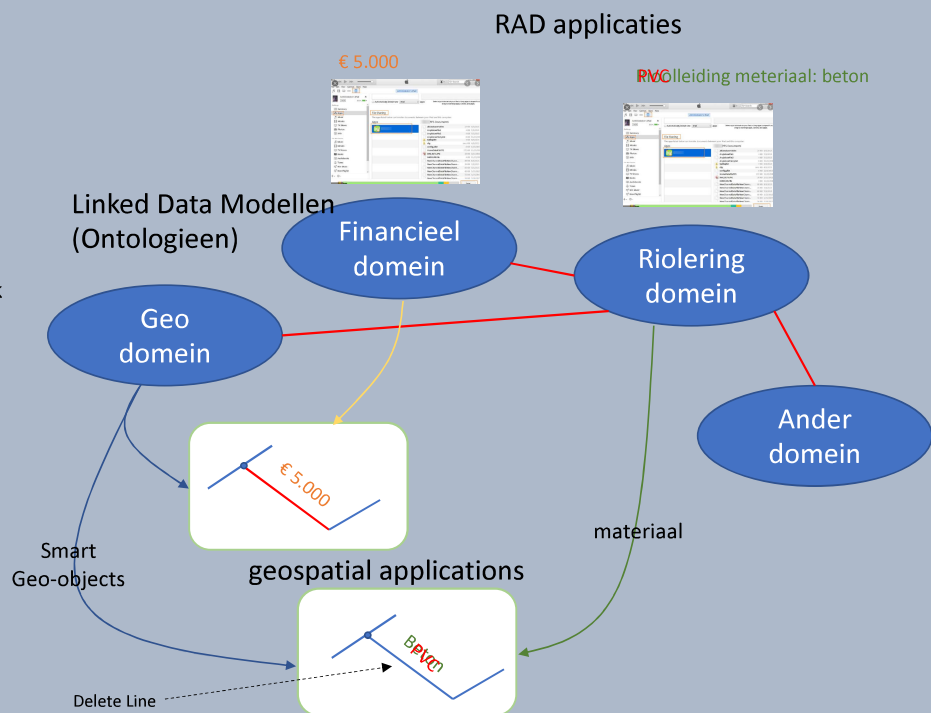


Linked Data en Smart – GEO (data kwaliteit)



Linked data en geospatial applicaties (RDF)

- Maak of gebruik bestaande ontologie
- Connecties over domeinen heen
- Maak eigen Geo model (ontologie)
 - Annotations en posities
 - Smart geospatial-objecten
 - Linken aan andere ontologieen
- Geospatial applicatie – enkelvoudige opslag meervoudig gebruik
- None Geo applicatie
- Realtime updates in Graph (triples)

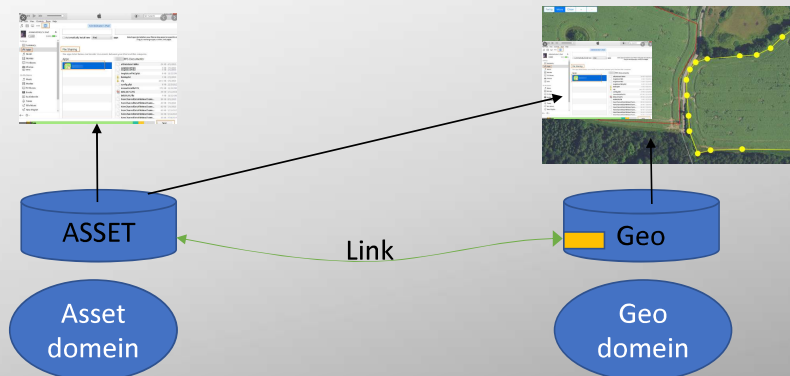


Enkelvoudige opslag meervoudig gebruik

Dus niet:



• Maar:



Summary

- Linked Data is zeer geschikt voor interoperabele infrastructuren en geospatial data.
- Linked Data is een alternatieve methode om relevante gegevens uit meerdere bronnen op te halen.
- Het is gebaseerd op de W3C(www)-standaard, gecombineerd met de OGC-standaard GeoSPARQL voor de behandeling van geospatial data.
- Het is een krachtige component om gegevensintegratie te maken op basis van de betekenis in plaats van exacte matching.
- Maakt geospatial gegevens FAIR (findability, accessibility, interoperability, and reusability)
- In tegenstelling tot een traditioneel datawarehouse worden alleen gegevens opgehaald die relevant zijn voor de uit te voeren query.
- Er bestaan een aantal implementatieplatforms voor Linked Data in combinatie met GeoSPARQL b.v. Virtuoso, Triply, Oracle



OSLO

Wat is OSLO?

- Open Standards for Linked Organisations – OSLO.
- De belangrijkste filosofie achter de OSLO-aanpak is het **waarborgen** van de **herbruikbaarheid** van bestaande internationale standaarden (W3C, EU ISA Core Vocabulaires, INSPIRE...) en het streven naar betrokkenheid van een groot aantal **verschillende belanghebbenden**, zoals de overheid, het bedrijfsleven en de academische wereld. Het standaardiseren van de betekenis van informatie is essentieel.
- Op dit moment bevat OSLO meer dan 18 domeinmodellen bestaande uit 1000+ definities gemaakt door meer dan 262 medewerkers van 107 organisaties.

Data opmaken volgens OSLO

OSLO implementatiemodel

OSLO Implementatieprofiel

= **Structuur + Context (betekenis)**

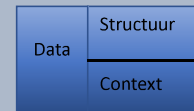
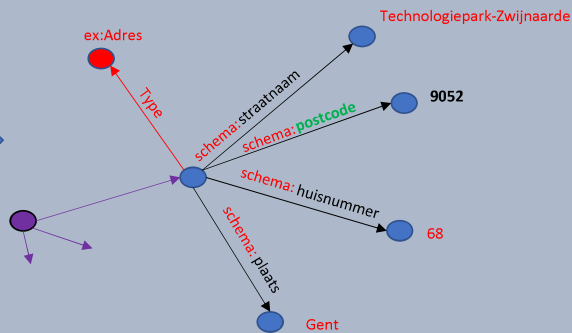
Context = betekenis

```
{  
  "@context": {  
    "schema": "http://schema.org/",  
    "ex": "http://example.org/"  
  }  
}
```

<http://schema.org/postcode>"9052"

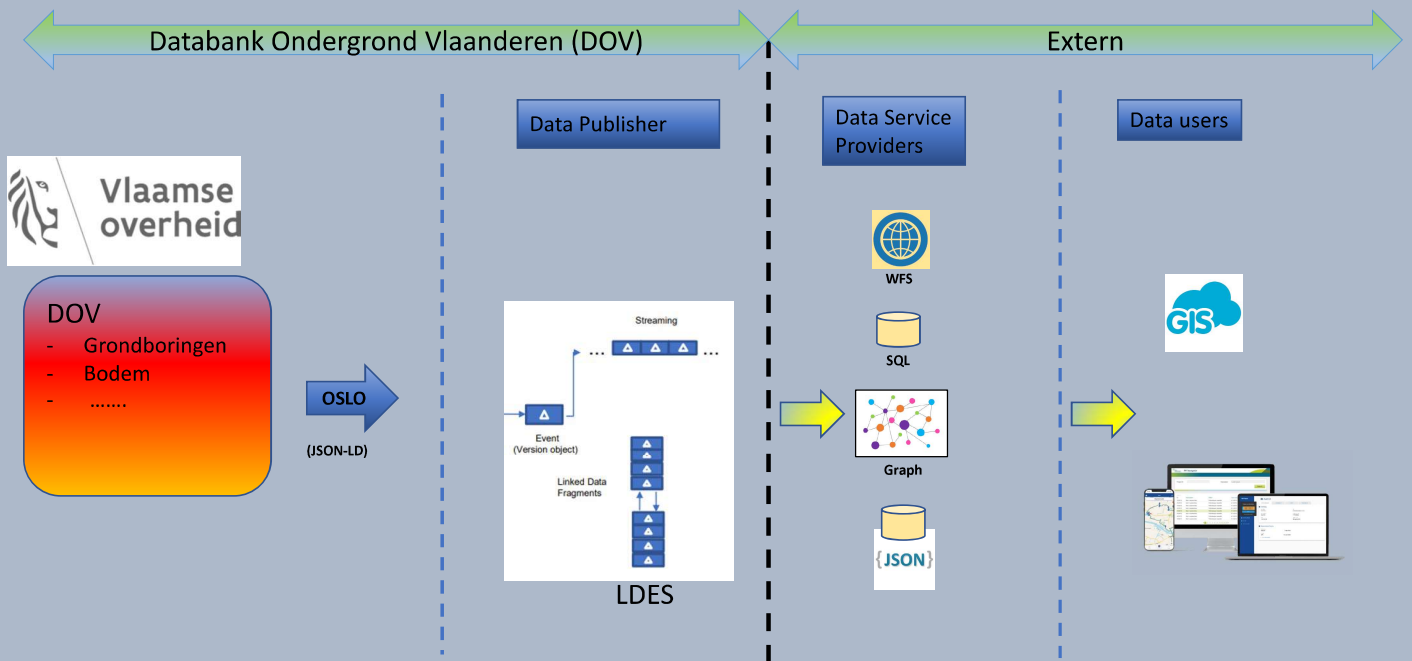
Data

Straat: Technologiepark-Zwijnaarde
Plaats: Gent
Huisnummer: 68
Postcode: 9052



OSLO
(JSON-LD)

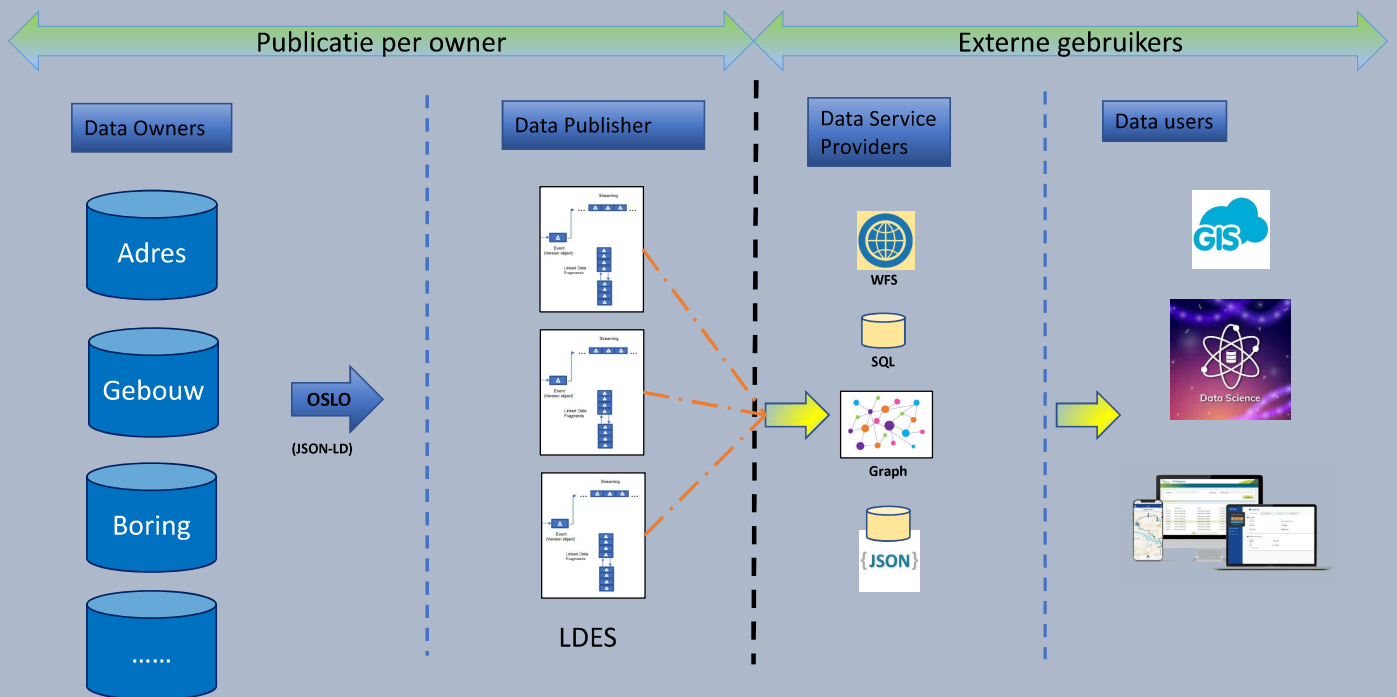
Publicatie van OSLO data



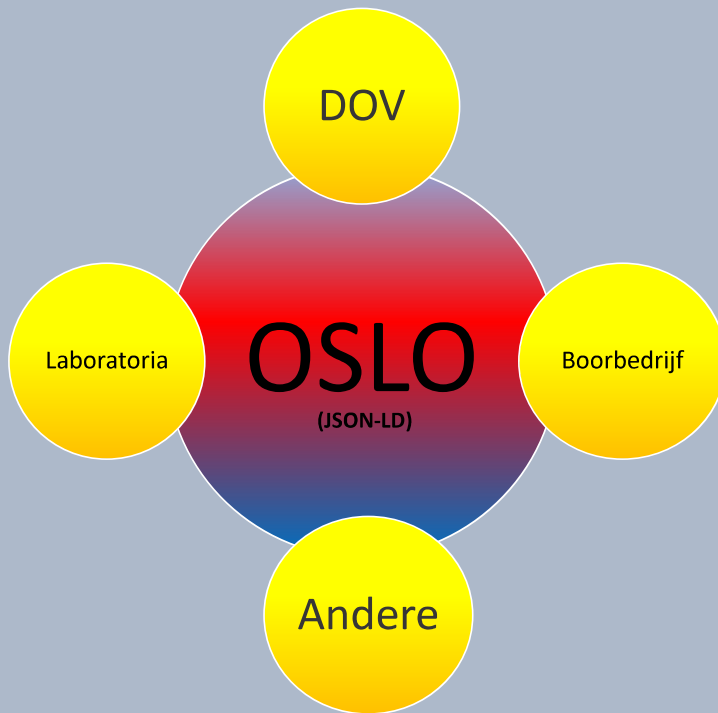
Enkele OSLO implementaties

- **OSLO Adres:** Dit model behandelt adressen en locatiegegevens, waardoor een gestandaardiseerde manier ontstaat om adresinformatie uit te wisselen.
- **OSLO Organisatie:** Dit model definieert eigenschappen en relaties met betrekking tot organisaties, zoals bedrijven, overheidsinstanties en non-profitorganisaties.
- **OSLO Persoon:** Het persoonlijke domeinmodel beschrijft kenmerken en relaties met betrekking tot individuen.
- **OSLO Grondgebied:** Het grondgebiedmodel omvat geografische gebieden, zoals gemeenten, provincies en landen.
- **OSLO Gebouw:** Dit model beschrijft eigenschappen van gebouwen en hun locatie.
- **OSLO Milieu:** Dit model behandelt milieuaspecten, zoals luchtkwaliteit, geluidsniveaus en bodemgesteldheid.

Publicatie van OSLO data



Data uitwisseling volgens OSLO



Samenvatting OSLO

- **Semantische interoperabiliteit:**
 - sector overschrijdend hergebruik van datasets.
 - naadloos integratie met andere datasets.

- **Harmonisatie van gegevens:**
 - consistente structuren en relaties van gegevens en diensten b.v. grondboringen.

- **Verbeterde vindbaarheid van gegevens**
 - makkelijker om relevante informatie te ontdekken en te gebruiken.
 - machine leesbaar.

- **Versnelde innovatie:**
 - ontwikkelaars, onderzoekers en bedrijven kunnen applicaties en diensten bouwen bovenop betrouwbare en interpreteerbare gegevensbronnen.



LIDIA

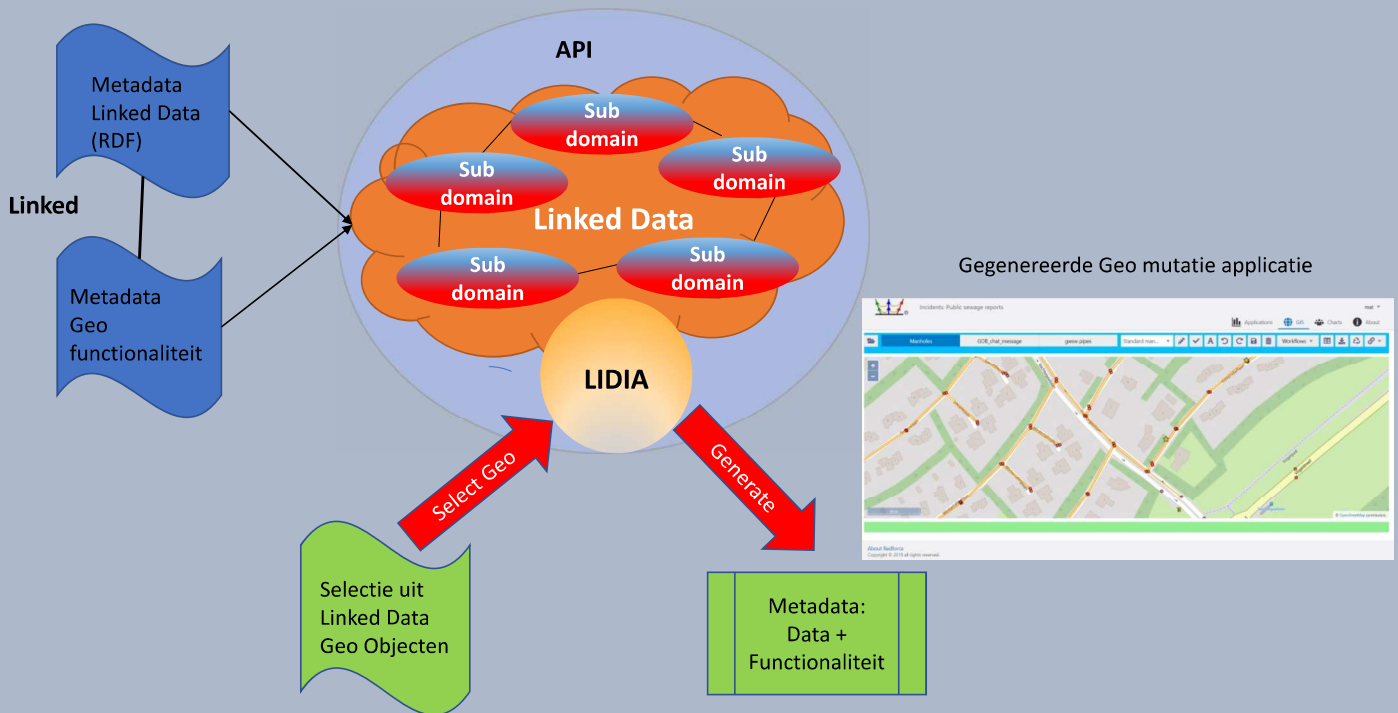
Waarom LIDIA

- **L**Inked **D**ata **I**ntegration **A**pplication – LIDIA.
- Onstaan door gebrek aan:
 - Integratie met bestaande tooling.
 - Integratie met andere data formaten.
 - GIS structuren (smart geo).

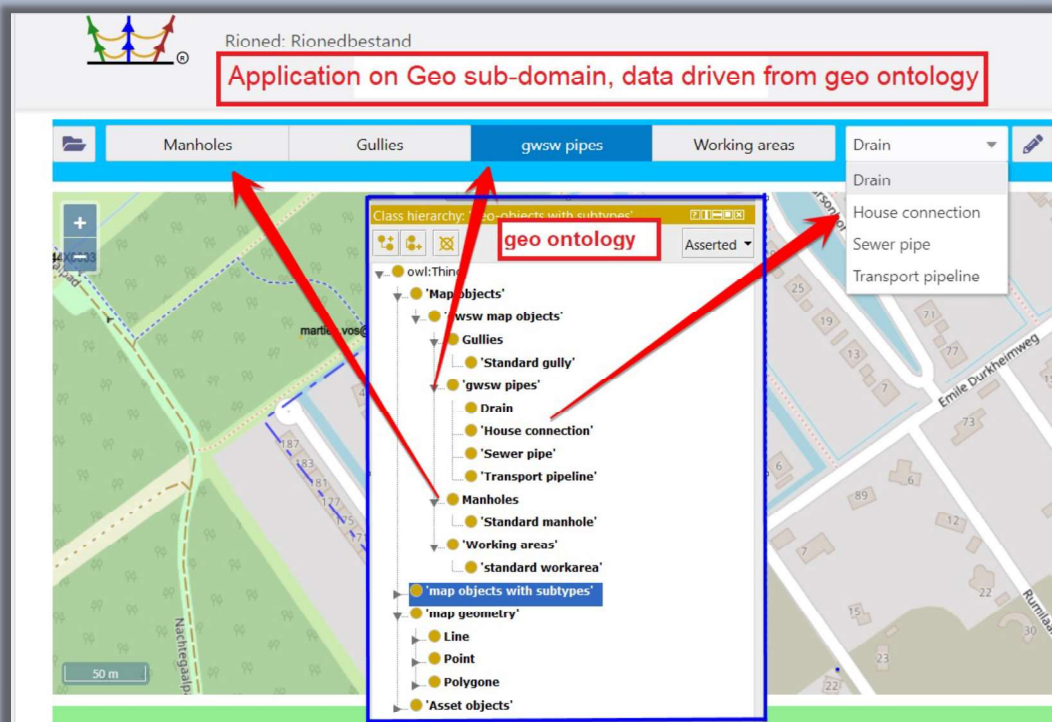
Waarom LIDIA?

- Toepassen Rapid Application Development (RAD) op LD. On demand applicaties.
- GIS toepassingen realiseren, optioneel:
 - live muteren van Linked data, **ook geometrie**.
 - Geo-mutatie applicaties **genereren**
- Betere datakwaliteit van geo data via smart geo objecten.
 - Mogelijkheid om te combineren met andere LD modellen.
- Toepassen van enkelvoudige opslag meervoudig gebruik.
- Combineren van LD met andere data structuren (Relationeel, Al..) in 1 query
- Ter beschikking stellen van Linked Data aan , 3th party's:
 - Analytische applicaties
 - Viewers (b.v. Geoserver)

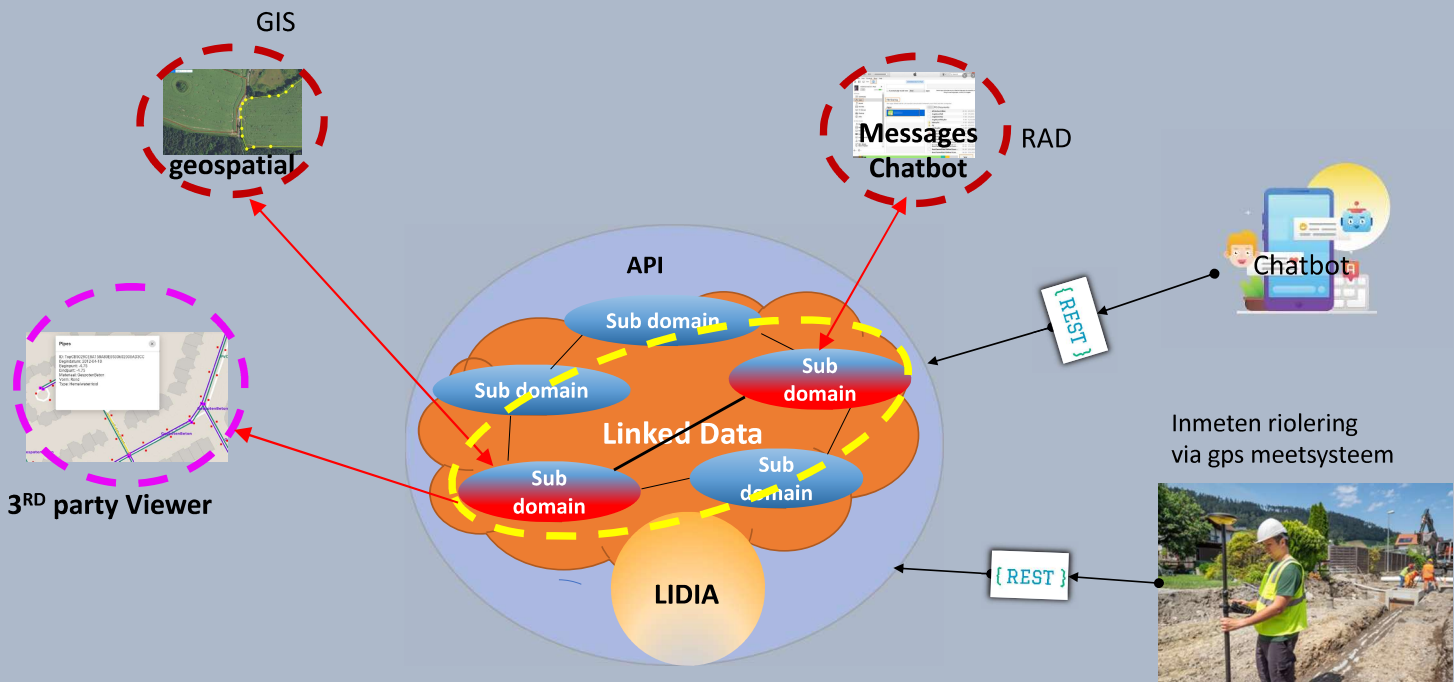
Genereren applicatie op basis van metadata in graph



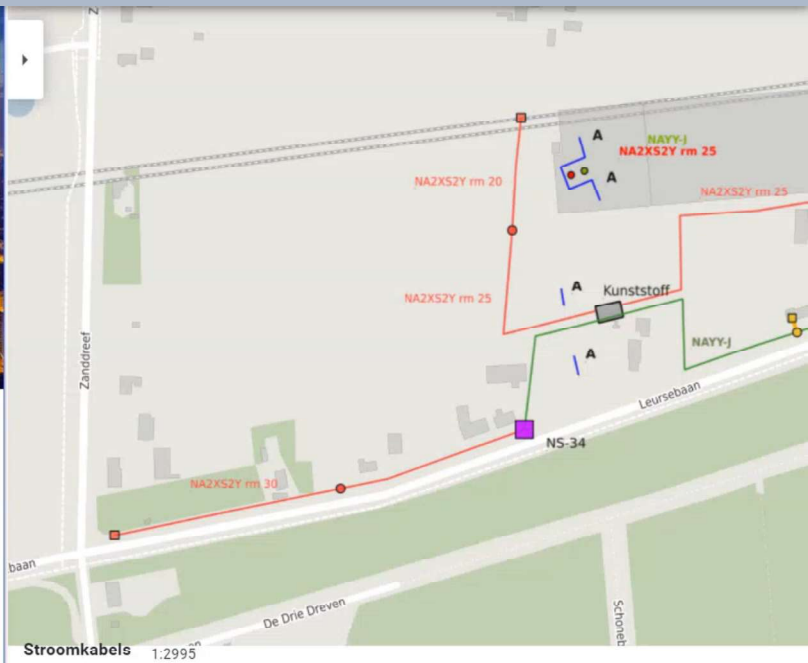
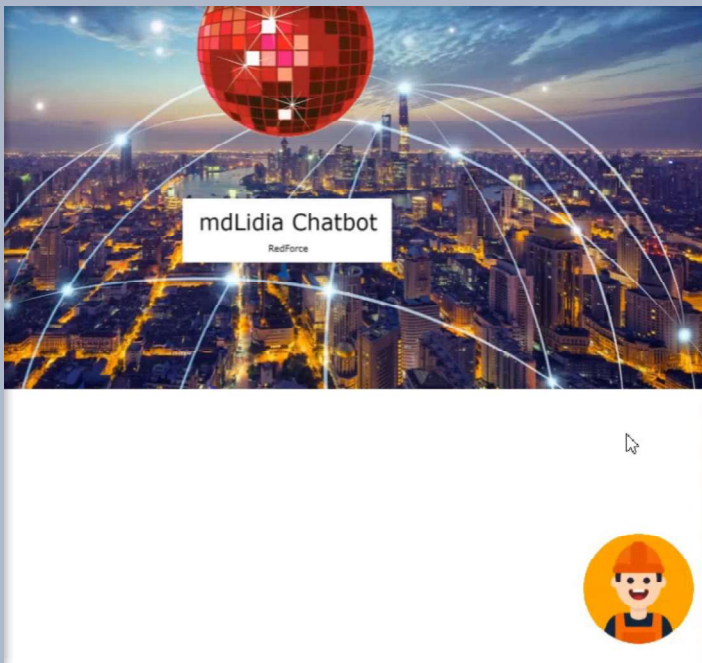
Geo-ontologie en data driven applicatie



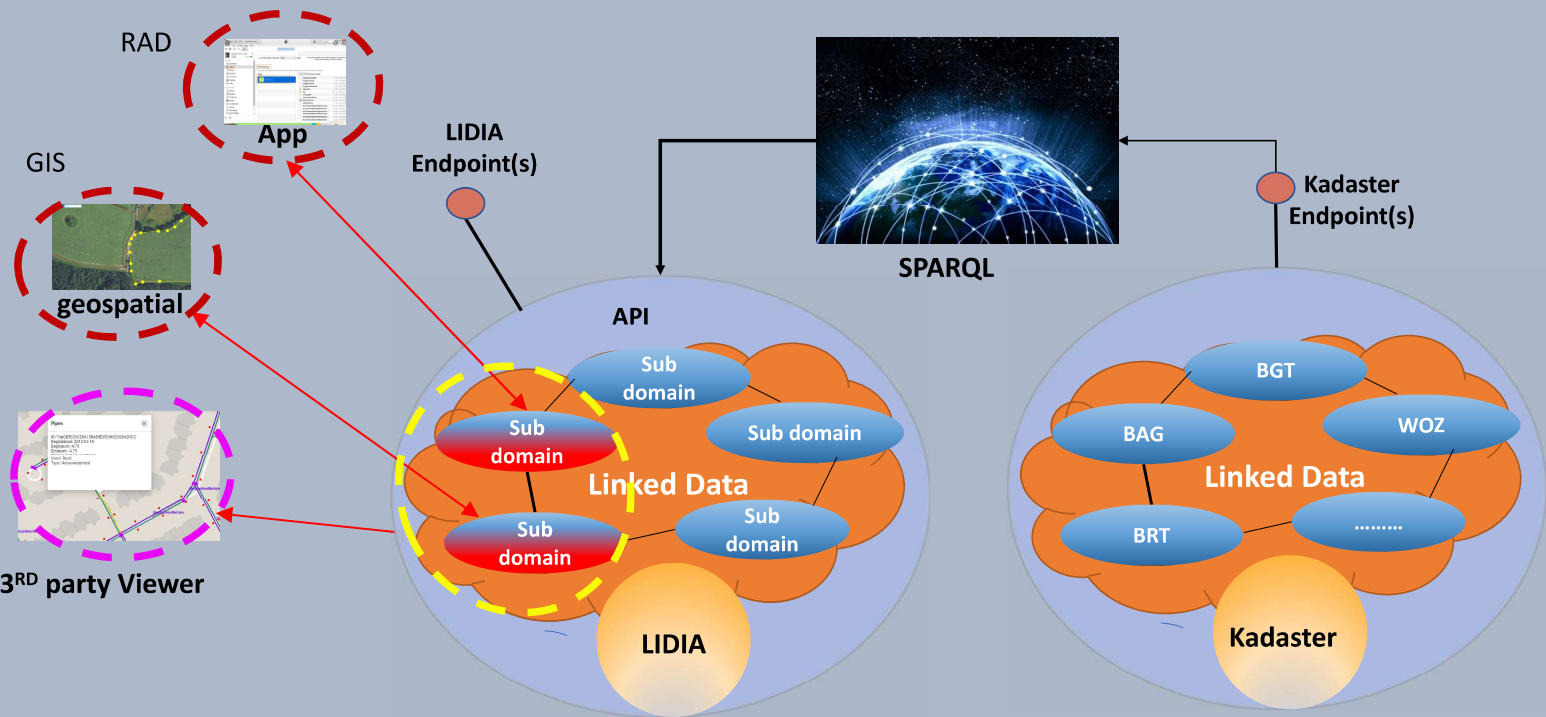
Voorbeeld LIDIA, RAD en REST



Voorbeeld LIDIA en schademelding (animatie)

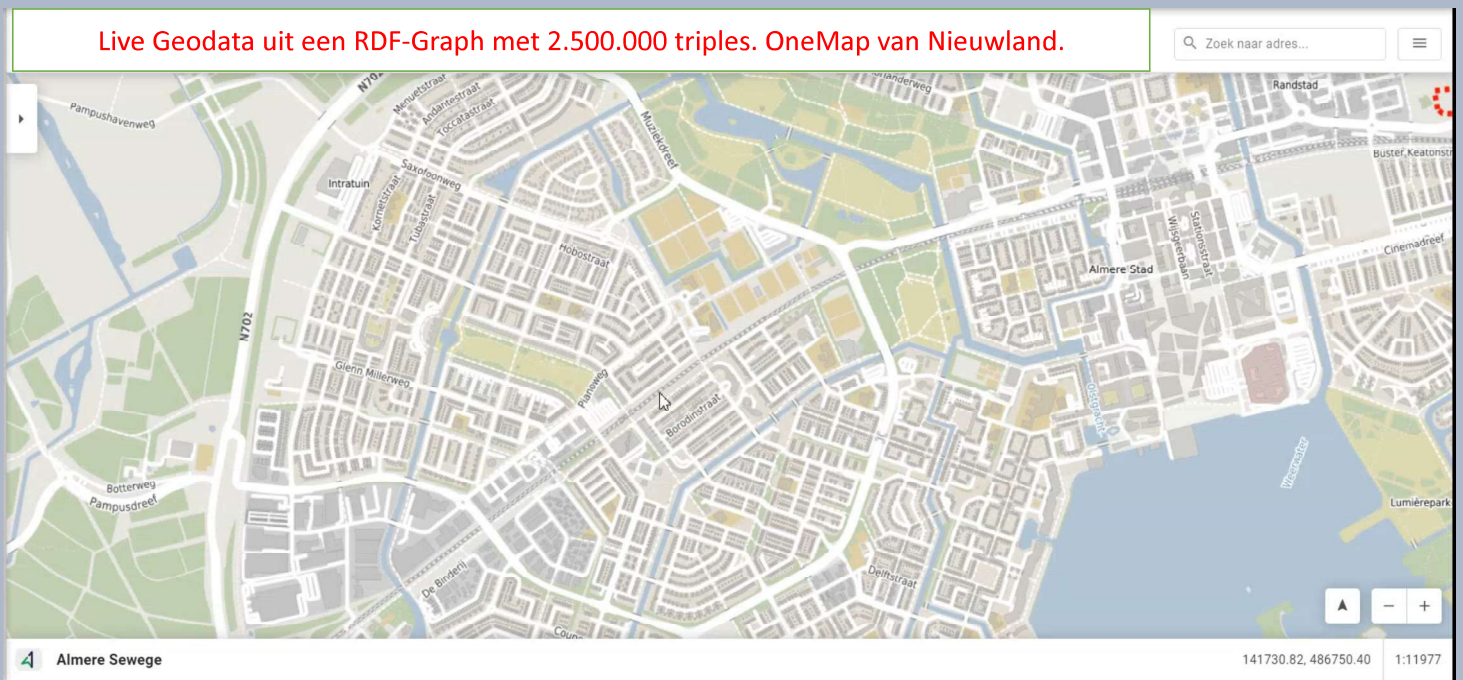


Voorbeeld LIDIA, RAD en federated queries



3th party Viewer integratie (animatie)

Live Geodata uit een RDF-Graph met 2.500.000 triples. OneMap van Nieuwland.



Overall Summary

- Linked Data produceren via 3 methoden:
 - Virtuele methode -> Virtuele laag van betekenis (semantiek) over bestaande data.
 - OSLO methode -> Data replicatie als Linked Data
 - Migratie methode -> Data migreren naar Linked Data
- Alle 3 de methoden kunnen gecombineerd worden.
- Geen probleem om Geo data op te nemen in deze methoden.
- LIDIA kan bij al deze methoden worden toegepast.
 - Linked Data
 - Relationaleel
 - JSON
 - Vector search
 - GIS .. en meer



Bedankt voor uw belangstelling

Vragen?

Martien Vos