

kadaster



De weg van Linked Data @ Kadaster

Erwin Folmer 02-07-2024
Erwin.Folmer@kadaster.nl

Kadaster Data Science Team

Wat doen we als Data Science Team:

AI (op aktes, op beeldmateriaal)

- Voorspellen bouwjaren
- Voorspellen verdiepingshoogtes
- Voorspellen graafschade
- Visualisatie – Dashboards - woningrealisatie
- Data analyse
- Zorgeloos Vastgoed – PODs
- Knowledge Graphs
- Duurzaam AI Bier: 't Perceeltje

Content

Context – Linked Data (@NL) – Kadaster

Why Linked Data @ Kadaster

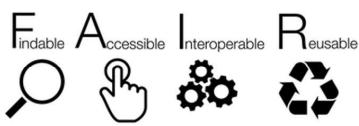
The road we took & lessons

Phase 1 – Kadaster Data Platform

Lessons & Renewed Focus

Phase 2 – The Kadaster Knowledge Graph

Latest Developments

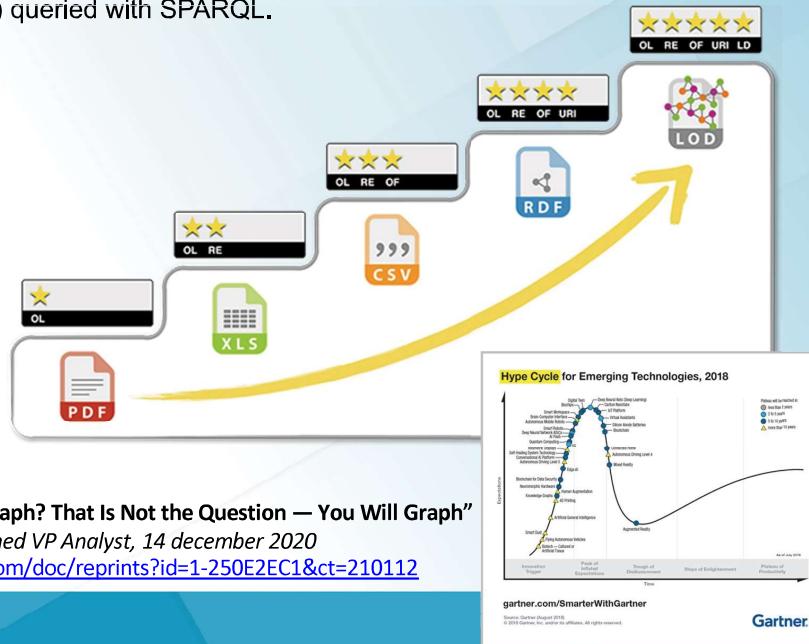


What is Linked Data?



Linked Data: A way of publishing data (reusable, web standards, semantics, related to open and big data). Data is stored as triples (RDF standard) and can be (federated) queried with SPARQL.

The screenshot shows a TED talk by Tim Berners-Lee titled 'The next web'. Below the video, there is a detailed graph visualization of linked data triples. Nodes include 'Elvira', 'Person', 'University of Twente', 'Organization', 'Erwin', 'Kadaster', 'House', 'Warnsveld', and 'Platform Linked Data Nederland'. Edges represent relationships like 'ReceivedPDFFrom', 'Type', 'WorksFor', 'Owns', 'idMarriedWith', 'isPartOf', and 'FoundedBy'. The graph is displayed on a white background with blue and black nodes and arrows.



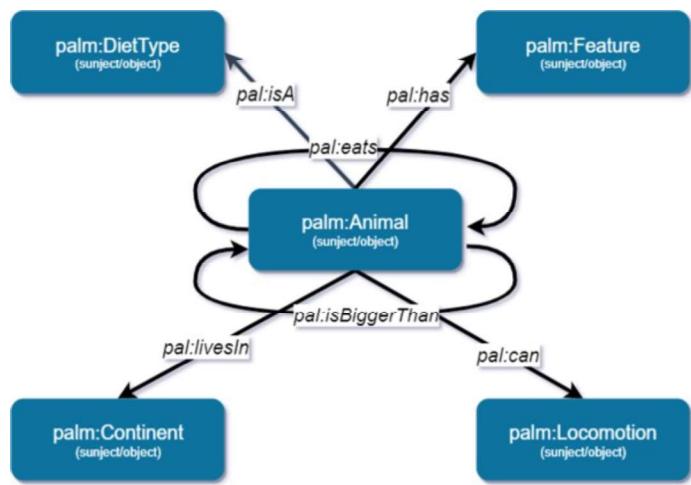
"To Graph or Not to Graph? That Is Not the Question — You Will Graph"
Mark Beyer, Distinguished VP Analyst, 14 december 2020
<https://www.gartner.com/doc/reprints?id=1-250E2EC1&ct=210112>

gartner.com/SmarterWithGartner
 Source: Gartner (August 2018)
 © 2018 Gartner, Inc. and/or its affiliates. All rights reserved.

Gartner

Linked Data in the Netherlands

- Platform Linked Data the Netherlands (PLDN)
- Billion Triples Cases



playalod / dieren

Search... CTRL | K LOGIN

Play-a-LOD

Browser

Table

SPARQL

Graphs 1

Services 1

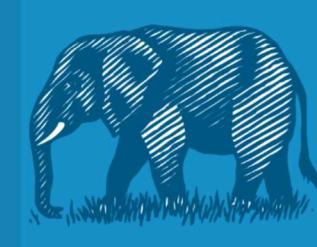
Assets 48

Insights >

<https://playalod.nl/dieren/id/Olifant>



Olifant



Dier

<https://playalod.nl/dieren/id/Olifant>

Type
Dier

Label
Olifant

SameAs
Elephant
Q7378

Depiction


Aantal kaarten
1

Heeft punten
3

Is SPO
S/O

Kadaster Labs | Grensreconstructie nl | Reisplanner | Plan je route | Kadaster Labs | verdiepingshoogte | https://data.osm.pldn.nl | OpenStreetMap - osm dataset

data.osm.pldn.nl

 Platform Linked Data Nederland

Innovatie door betekenisvol verbinden

Home SPARQL Endpoint OSM Dataset PLDN Website

OpenStreetMap (OSM) Dataset



De OpenStreetMap dataset is een publicatie van de Linked Data versie van de welbekende geospatiale datasets van OpenStreetMap. De beschikbaarstelling van deze dataset is een samenwerkingsverband tussen de volgende drie partijen:

- Platform Linked Data Nederland (PLDN)
- CGI Infra and asset management lab
- Triply

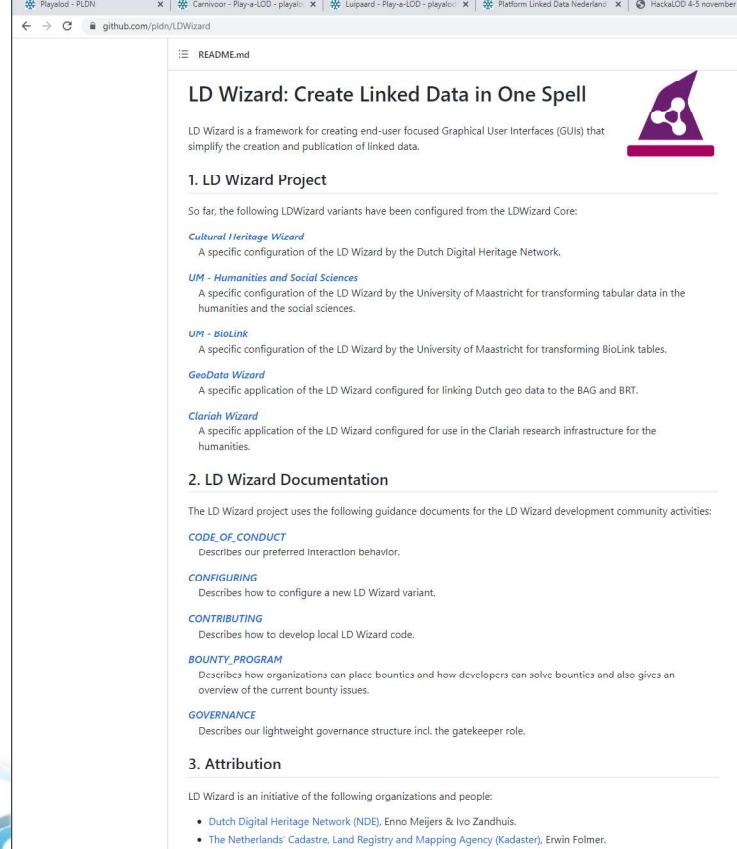
OpenStreetMap (OSM)

OpenStreetMap (OSM) is een project dat als doel heeft om vrij beschikbare en bewerkbare geografische gegevens te verzamelen, zodat daaruit landkaarten en andere diensten kunnen gemaakt worden. Enieder die zich inschrijft, kan eraan meewerken. Het invoeren en aanpassen van de geografische data steunt volledig op vrijwilligers. Het is mede opgericht uit onvrede met de hoge kosten die commerciële kaartbedrijven vragen voor hun producten. Reeds meer dan 20 miljoen kilometer wegen, fiets- en voetpaden werden in kaart gebracht door middel van het overnemen van gegevens van luchtfoto's en zelf op pad te gaan met een GPS-apparaat. In juni 2020 waren elke dag gemiddeld zo'n 6300 vrijwilligers actief aan het bijdragen.

Bron: OSM Wiki

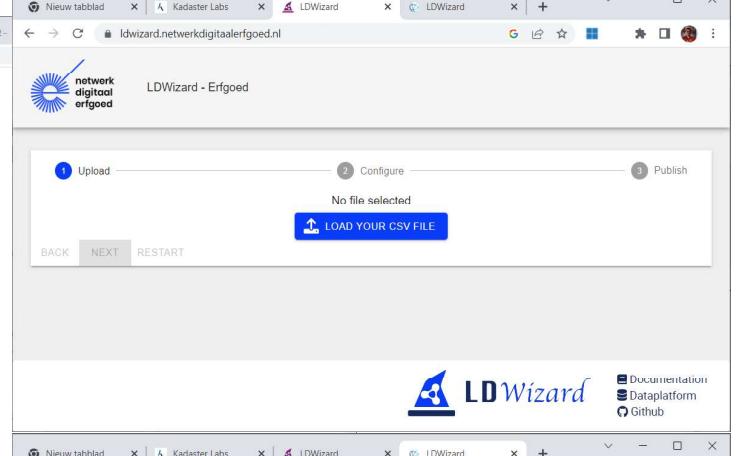
welke_straten_heeft...csv ▾ Alles bekijken x

github.com/pldn/LDWizard



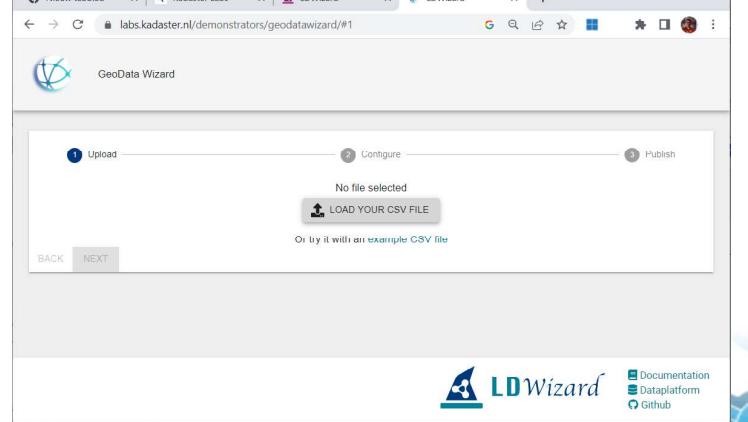
The screenshot shows the GitHub page for the LD Wizard project. The README.md file contains information about the project, including its purpose (creating end-user focused Graphical User Interfaces for linked data), variants (LD Wizard Project, Cultural Heritage Wizard, UM - Humanities and Social Sciences, UM - BioLink, GeoData Wizard, Clariah Wizard), and documentation links.

ldwizard.netwerkdigitaalergoed.nl



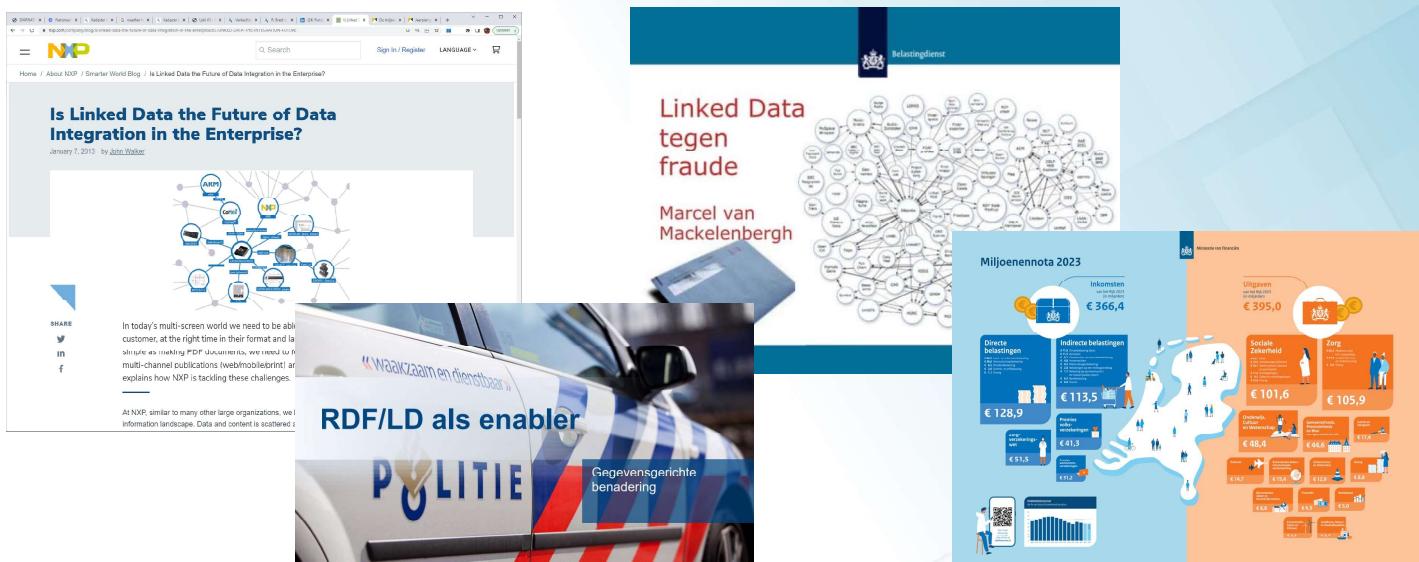
The screenshot shows the LD Wizard interface for the Erfgoed variant. It features a header with the networkdigitaalergoed logo and the title "LDWizard - Erfgoed". Below the header is a navigation bar with three steps: Upload, Configure, and Publish. A "LOAD YOUR CSV FILE" button is prominently displayed. At the bottom, there are "BACK", "NEXT", and "RESTART" buttons.

labs.kadaster.nl/demonstrators/geodatwizard/#1



The screenshot shows the LD Wizard interface for the Kadaster variant. It features a header with the Geodata Wizard logo and the title "GeoData Wizard". Below the header is a navigation bar with three steps: Upload, Configure, and Publish. A "LOAD YOUR CSV FILE" button is prominently displayed. At the bottom, there are "BACK", "NEXT", and "TRY IT WITH AN EXAMPLE CSV FILE" buttons.

The Billion Triples Cases



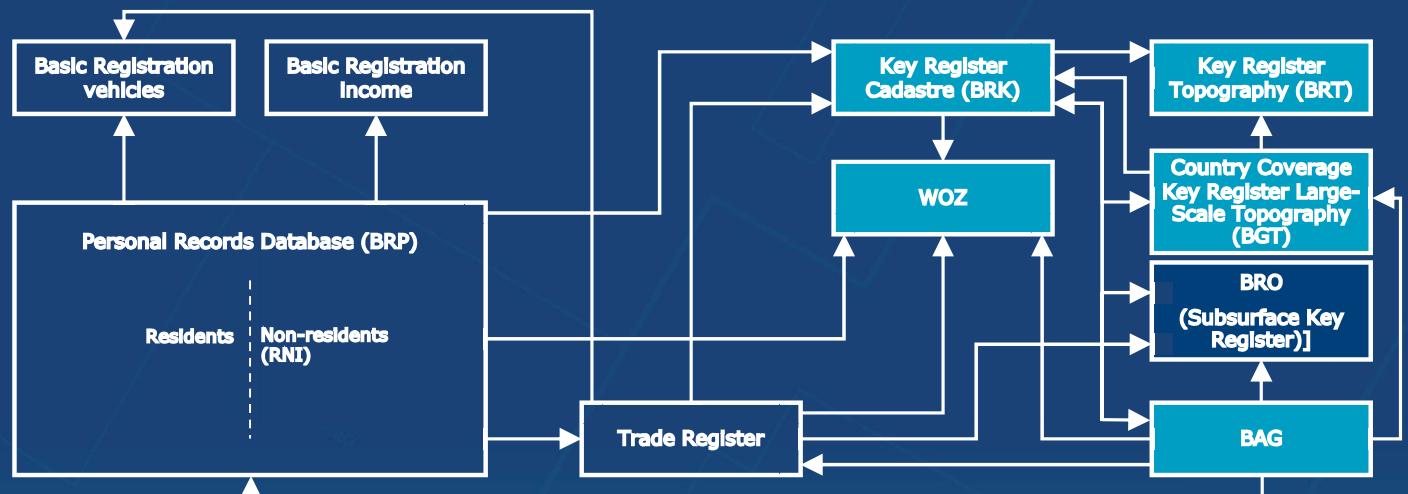
<https://www.linkedin.com/pulse/fundamentals-digitalisation-towards-holographic-flores-bakker/>

Do you know Kadaster?

Data per Key Register



Key Registers in the Netherlands



Key registers and national facilities

Key Registers

National facilities

Other Tasks

Key Register Cadastre (BRK)

Key Register Addresses & Buildings (BAG)

KLIC (Dutch Cable and Pipeline Information Centre)

- Key Register Cadastre (BRK) Registration System
 - Registrations spread across government organizations.

Kadaster source holder of:
Key Register Topography (BRT)

Key Register Large-Scale Topography (BGT)

Schematic view

- Key Register Cadastre (BRK)
 - Key Register Topography (BRT)

- 5 registrations from other organisations (national facilities)
Valuation of real estate (WOZ)

Energy Labels

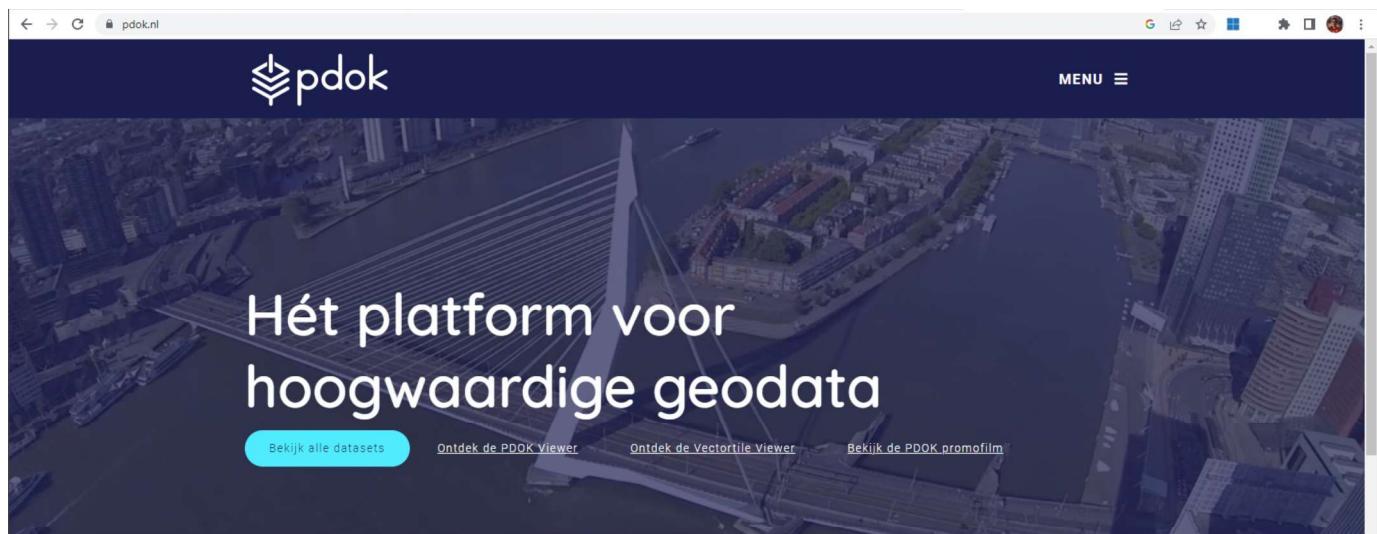
Act on the Recognisability of Public Law Restrictions and Real Estate (WKPB)

National Triangulation Network

Ruimtelijkeplannen.nl

Defense

• — •



Waarom PDOK?

1

Kwaliteit

PDOK levert hoge kwaliteit (standaard ISO 25010)

Hoogwaardige geodata

PDOK biedt hoogwaardige en actuele open geodatasets

Webservices / API's

De geodata is benaderbaar via standaard OGC- en W3C

ster



Key figures 2022

465.800

Deeds of transfer

494.500

Deeds of mortgage

91.100

Survey entries

29.1 billion

PDOK views

1.2 billion

BGT views

1.008.000

KLIC notifications

26.4 million

Information Products

3.5 billion

BAG Inquiries

49.1172



Summary

Kadaster = data



kadaster

Kadaster Data Platform

The development of the Spatial
Data Platform of the Future

Why?

kadaster



vision Kadaster

Making reliable data accessible to everyone. Through a single digital gateway. That is our job, by connecting and providing certainty. When data from different sources is combined, valuable insight is generated offering solutions for challenges in society. Our contribution to a sustainable and livable society. In the Netherlands and abroad.

Kadaster's Ambitions



Providing certainty of ownership and use of anything above and below the surface



Offering a platform for anyone to handle spatial information anywhere and anytime



Supporting society in using spatial information to solve issues that matter



What is the problem with copying?

Not only intrinsic quality

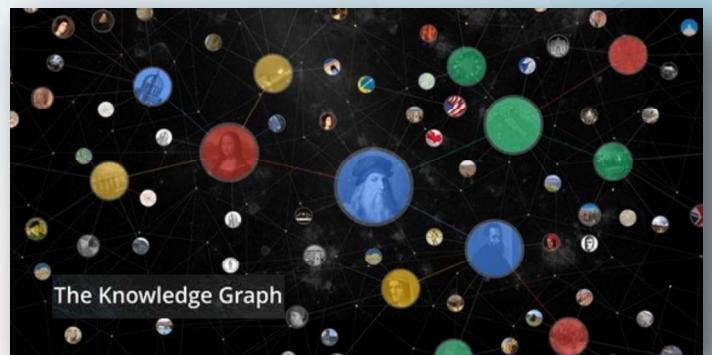
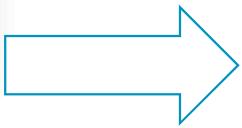
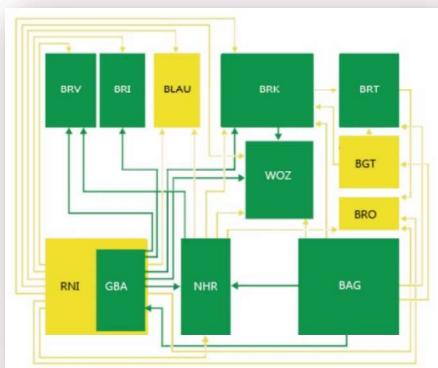
- The leaflet (metadata) is essential
- Semantics
- Provenance
- Transparency is key!



What is the problem with not using metadata?

- Size: 100
- Geometries

From data silo's to connected information



closed

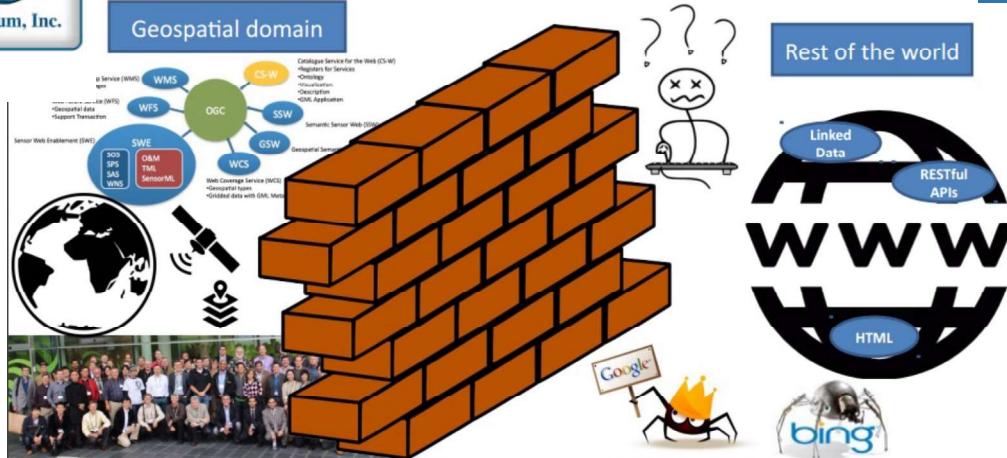


open



Spatial data: to the web

Reaching more people with spatial data



What is the problem?



hofstraat 110



Press F11 to exit full screen

All

Maps

Images

Shopping

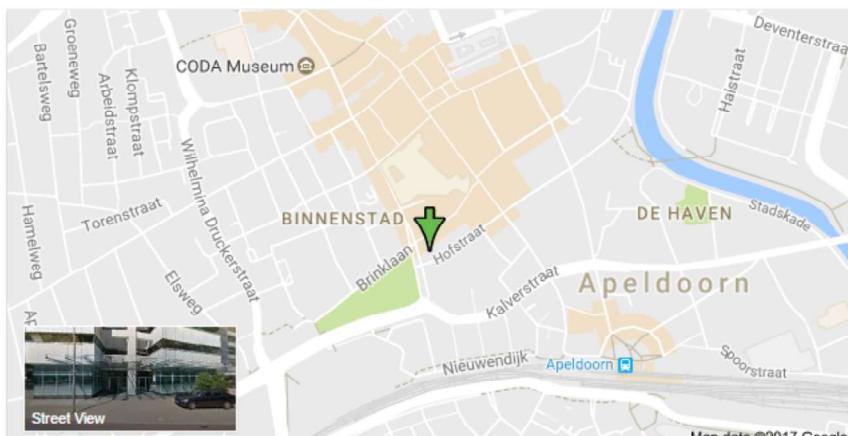
Videos

More

Settings

Tools

About 1.230.000 results (0,71 seconds)



Hofstraat 110, 7311 KZ Apeldoorn

[Get directions](#)

Kadasterkantoren

<https://www.kadaster.nl/kadasterkantoren> ▾ [Translate this page](#)

Hofstraat 110 7311 KZ Apeldoorn Telefoon receptie: 088-183 20 00. Routebeschrijving kantoor de Grift.

Particulier · Woningwaarde · Eigendom · Grenzen ...

Analytics Friendly

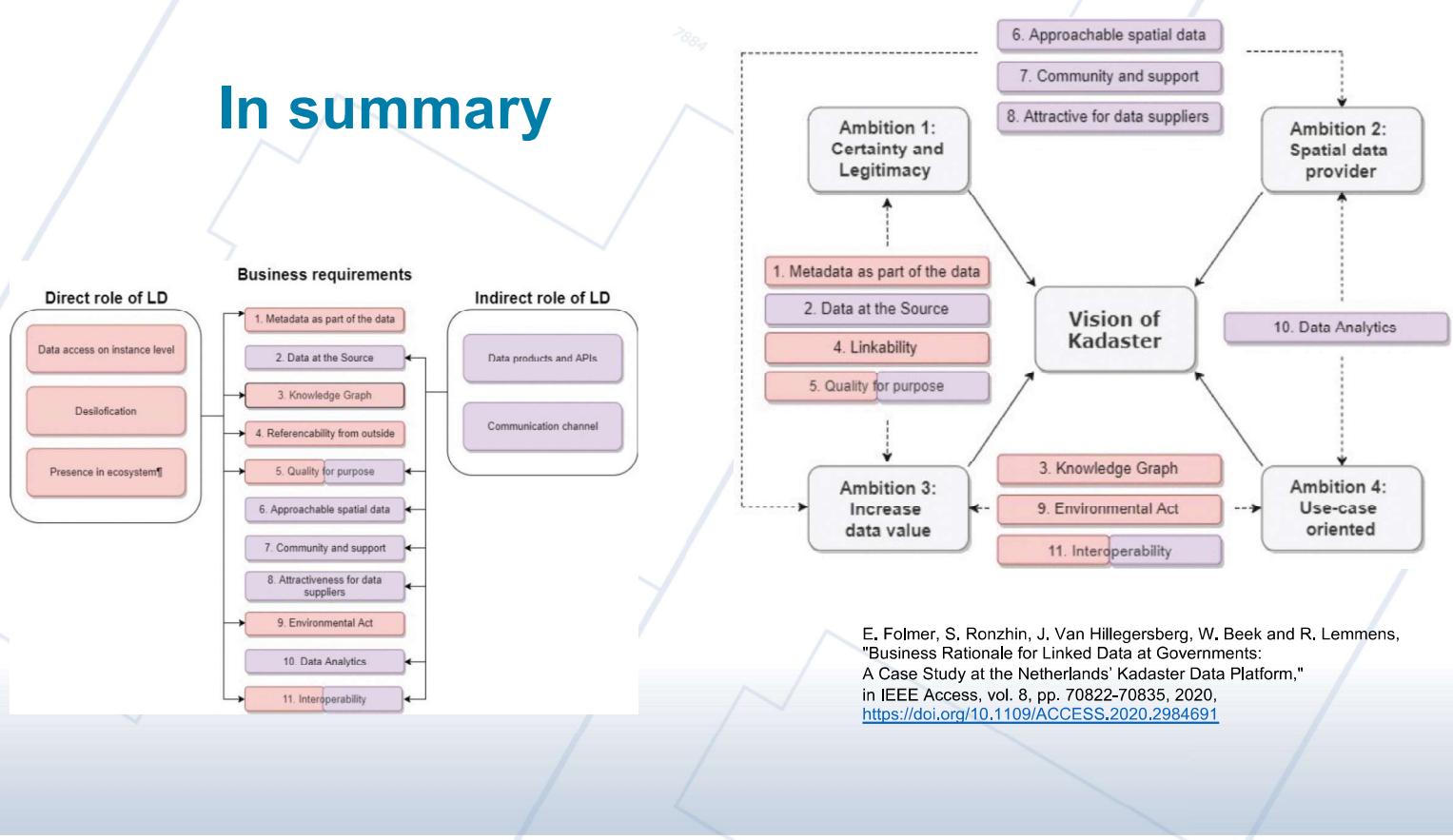
(Self Service GIS)

- Citizen (non GIS professional)
- Data at the Source – Data Ecosystem
- Semantics (Transparency, Reproducability)
- The Web (only a browser)
- Map integration – GIS functionality
- (integration with more advanced tools)

(GIS = Geographic Information System)



In summary





kadaster

Kadaster Data Platform

The development of the Spatial
Data Platform of the Future

Phase 1 - KDP



kadaster

Demo

A dataset as linked data:
<https://data.labs.kadaster.nl>

Live

A Dataset as linked data:

<https://data.labs.kadaster.nl>

(go to top10nl)

A sparql query on BRT (e.g. replace bunker with universities or ziekenhuizen)

Show uri derefencing....Trifid: from Zazuko.

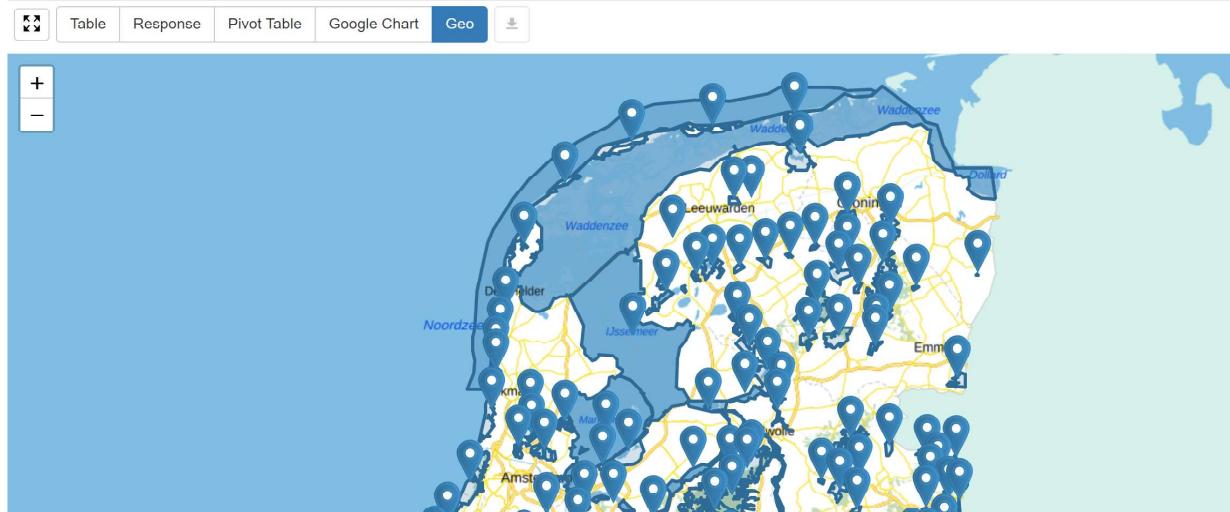
<https://brt.basisregistraties.overheid.nl/brt/id/registratie/NL.TOP10NL.100099668>

Data story based on sparql queries

(religieuze gebouwen)

https://data.pdok.nl/sparql

```
1 v prefix brt: <http://brt.basisregistraties.overheid.nl/def/top10nl#>
2 v prefix geo: <http://www.opengis.net/ont/geosparql#>
3 v select ?shape ?shapeLabel {
4 v   [ a brt:Natuurgebied;
5 v     brt:naamNL ?shapeLabel;
6 v     brt:soortnaam "natura2000-gebied"@nl;
7 v     geo:hasGeometry/geo:asWKT ?shape ].
8 }
9 limit 1000
10
```



Secure | https://www.edparsons.com/2017/09/cant-link-exist/

My Blog Events & Press

Rather than publishing online a database of railway station locations in the Netherlands and expecting a user to then query the database for "Amsterdam Centraal Station", publish the database giving each record a URL so for example Amsterdam Centraal Station becomes;

<https://brt.basisregistraties.overheid.nl/top10nl/id/gebouw/10262520>

Now this is something I can paste into an email, tweet or even share on Facebook !

Kudos to the Dutch Kadaster for taking this approach and providing this example, Ordnance Survey you could do the same ?

This approach also results in such data becoming part of the "mainstream" web indexable and searchable, but I argue the key benefit is the "linkability"

The Spatial Data on the Web best practice document, something of course I recommend you taking a longer look at provides many

Ed Parsons Retweeted

Sam Zipper @ZipperSam

Google Earth Engine + @NASA_Landsat = annual 30 m resolution #irrigation maps!

Cool new study by @JillDeines: onlinelibrary.wiley.com/doi/10.1002/20...

Annual irrigation dynamics in t... Sustainable management of agricultural water resources onlinelibrary.wiley.com

Sep 12, 2017

Sep 11, 2017

Embed View on Twitter

Lessons learned

It can be much less work than expected, but in the end it is not.

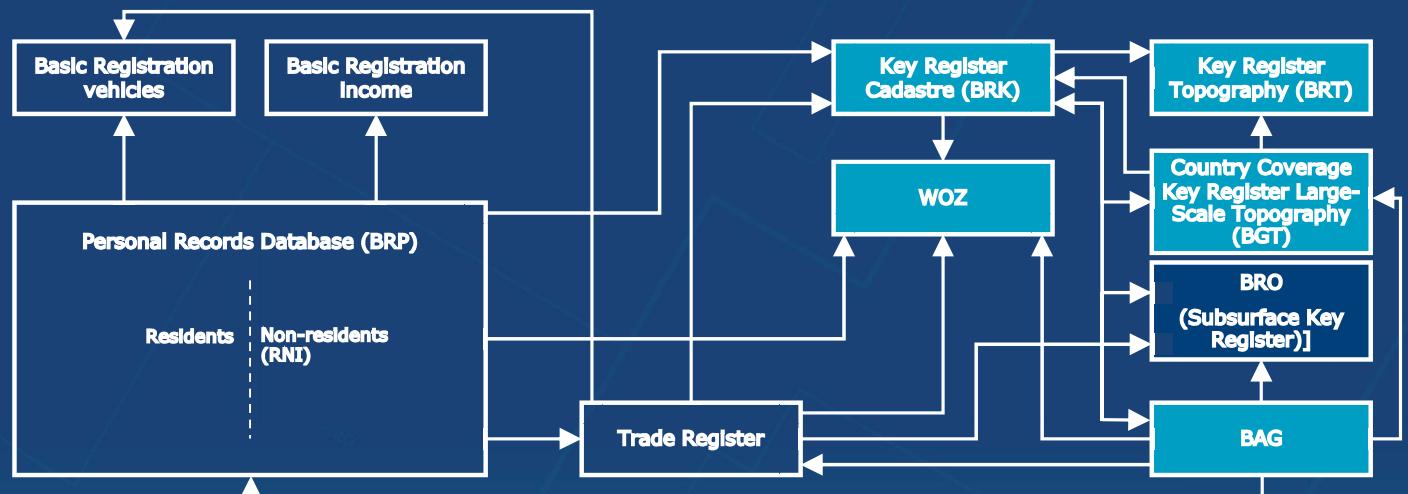
Projects are interfered with other issues:

- Improvements in data models (the “ultimate” model)
- Data quality issues (completeness, metadata, etc)

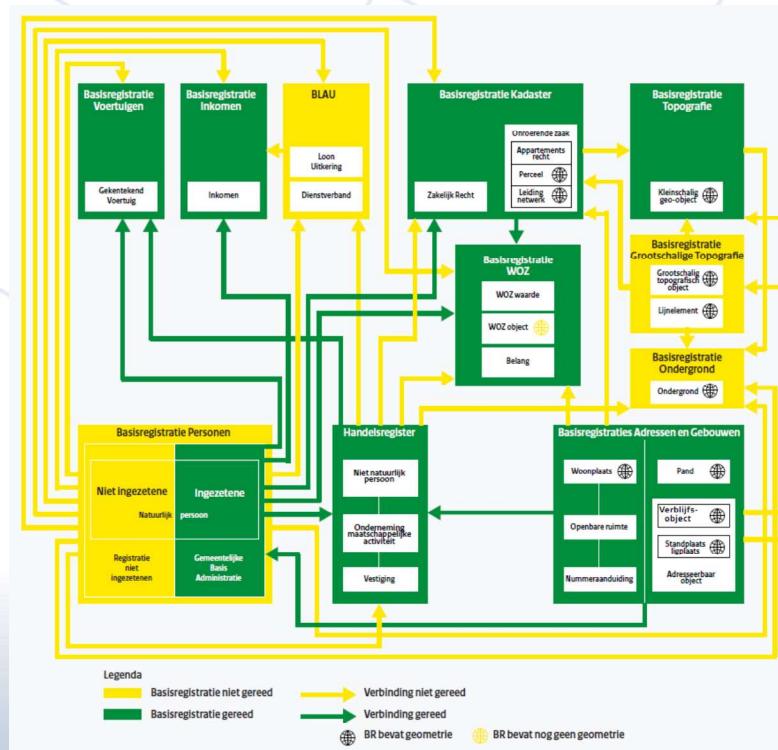
My advise: Keep it simple

- Sparql queries complex
- Federation....ahum...
- GeoSparql partly supported
- 4 star Linked Data

key registers in the Netherlands



Dutch Humour: System of Silos





Conclusion (~2019):

**We created the ability...
...but have no links...still data silo's.**



kadaster

Kadaster Knowledge Graph

Focus on Integrating Data for
Analysis (the Unknown Question)

Phase 2
Experiments 2020-2021
Production & Usability 2022-2023

kadaster

Can you imagine?

- That in 2023 we still offer data silo's?
- That citizens can only ask through Wikipedia or Google questions like; what are the churches built before 1900? And based on authentic government data no direct answer is possible?



Geo-informatie voor iedereen

What is our ambition?

(<https://www.kadaster.nl/over-ons/beleid/meerjarenbeleidpt>)

Wij vinden dat geo-informatie voor iedereen toegankelijk moet zijn. We bouwen ons platform verder uit met mogelijkheden voor analyse van data. Daarnaast gaan we datasets onderling koppelen. Voor de ontwikkeling ervan gebruiken we concrete vraagstukken van gebruikers.



vision Kadaster

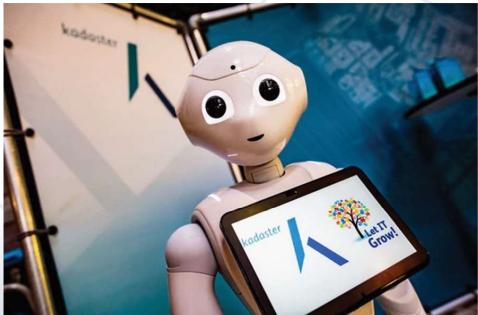
Making reliable data accessible to everyone. Through a single digital gateway.
That is our job, by connecting and providing certainty.

When data from different sources is combined, valuable insight is generated offering solutions for challenges in society. Our contribution to a sustainable and livable society. In the Netherlands and abroad.

My dream!



I search for a palace (BRT) with surface of >1500m² (BAG) which is a national monument (RCE), and has a value of >1.000.000 (WOZ) in the neighborhood called Berg en Bos (CBS).



Paleis het Loo

<https://brt.basisregistraties.overheid.nl/top10nl/id/gebouw/101895841>

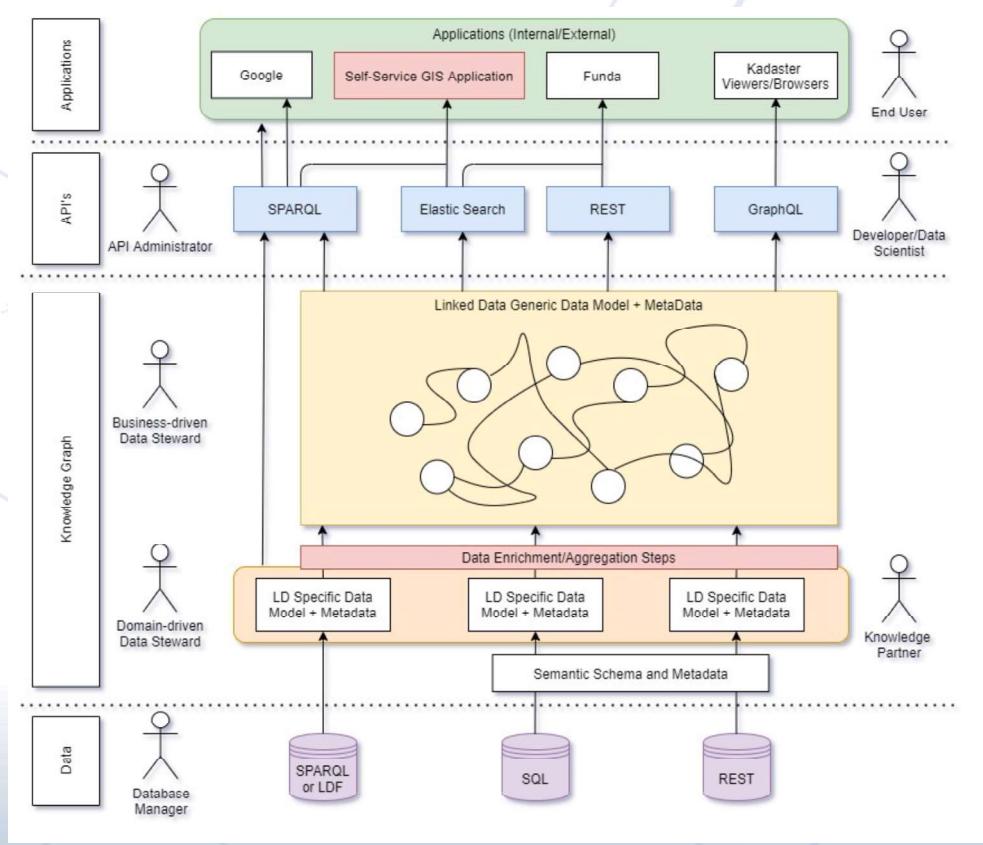
Essence of Kadaster KG project

It is not new, it was already possible.....but expensive, not scaleable, not open, etc...

The main differences are:

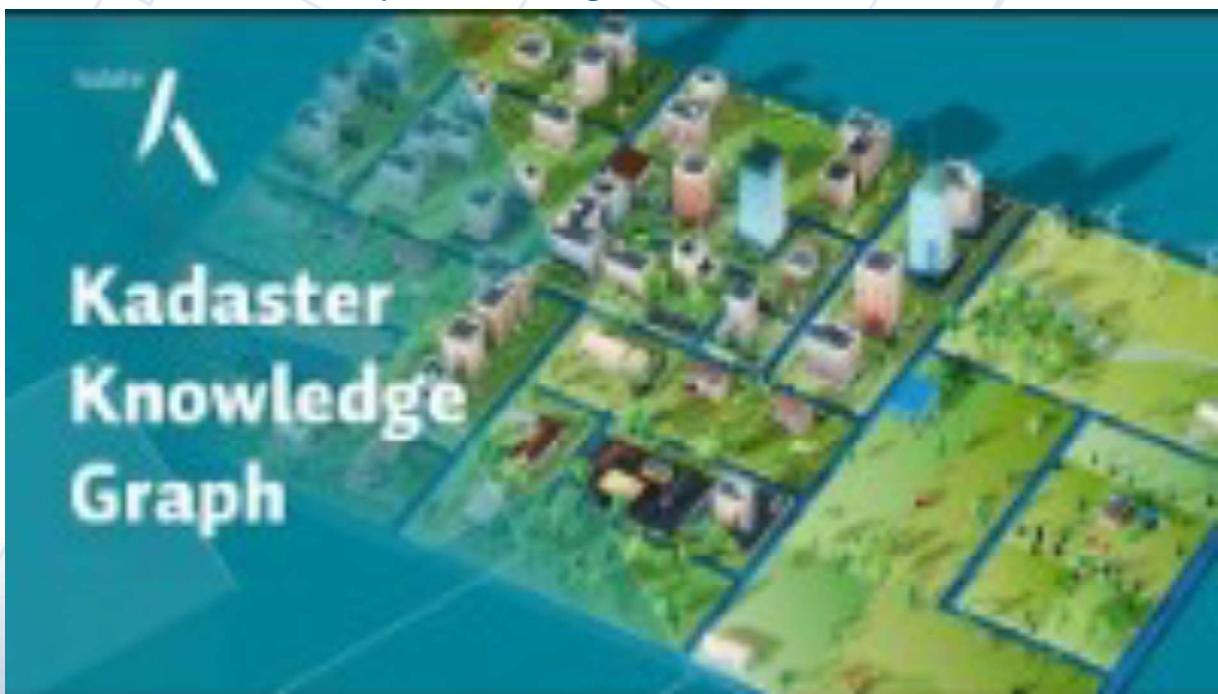
- Web standards based (linked data, interoperable)
- No spatial links, but administrative links.
- Data at the Source (in demo: 6 sparql endpoints federation)
- Application independent.

The Architecture for the KKG



Kadaster Knowledge Graph

<https://data.kkg.kadaster.nl>



Kadaster Knowledge Graph

Live demo

Movie? – screenshot op deze site

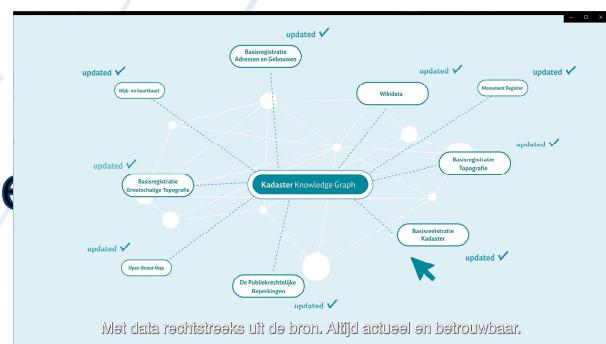
Data.kkg.kadaster.nl

Query: niet; door naar data story

Data Story: algemene queries voor kkg
gebruik

Data model: time for some art.

Production ready!



Lessons Learned

- It works for integrated data analysis and it is being used
- Promoting the KKG (not datasets as LD)
- One KKG for World Peace? Or Many?
 - Complexity for users
 - Complexity for tools
 - Maintenance
- Make it easier!
 - Educate
 - Inspire
 - Tools



kadaster

Kadaster Knowledge Graph (KKG)

Latest Developments on Usability

- Inspiration
- Easy/less SPARQL

Warning: Work in Progress!

kadaster

<https://labs.kadaster.nl/cases/ar-applicatie>

The screenshot shows a web-based API interface. At the top, there is a red box highlighting the URL: "GET: https://api.labs.kadaster.nl/queries/dst/ar-demo-data-voor-plaats/run?". Below this is a "Variables" section with a red box around the variable "vbold" set to "0003010000129471". A "View populated query" link is also present. The main area displays a large amount of RDF query code. To the right of the code is a mobile application interface showing building information for a specific address.

Variables

vbold
0003010000129471

View populated query

```
1 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
2 PREFIX dct: <http://purl.org/dc/terms/>
3 PREFIX sdo: <https://schema.org/>
4 PREFIX brt_def: <http://brt.basisregistraties.overheid.nl/def/to>
5 PREFIX kad: <https://data.kkg.kadaster.nl/kad/model/def/>
6 PREFIX geo: <http://www.opengis.net/ont/geosparql#>
7 PREFIX sdob: <http://schema.org/> #error in CBS data
8 PREFIX sor: <https://data.kkg.kadaster.nl/sor/model/def/>
9 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
10 prefix time: <http://www.w3.org/2006/time#>
11 PREFIX wbk: <https://data.labs.kadaster.nl/cbs/wbk/vocab/>
12 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
13 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
14- SELECT DISTINCT ?gebouwNaam ?buurtNaam ?buurtID ?gemeenteNaam ?gebouwjaar ?gebruiksdoel ?basischoolafstand ?reinstandafstand
15 bind(iri(concat('https://data.kkg.kadaster.nl/id/verblijfsobject',
16 ?vbo
17 a sor:Verblijfsobject;
18 sor:gebruiksdoel/skos:prefLabel ?gebruiksdoel;
19 sor:maaktDeelUitvan ?gebouw;
20 sor:hoofdadres ?nummeraanduiding;
21 sor:oppervlakte ?vloerooppervlakte .
22 ?gebouw
23 a sor:gebouw;
```

Adres
't Spiker 24
7231JT

woonfunctie

Bouwjaar:	1976
Vloer Oppervlakte:	394 m ²
Perceel Oppervlakte:	530 m ²
Gemeente:	Zutphen
Buurt:	Dorp Warnsveld
Aantal basisscholen binnen 3km:	12
Afstand treinstation	2.7 km

Meer informatie

Klopt er iets niet?

Feature 3: Building Information

- Selected building identifier and building part identifier is used as input, all attribute information for both objects are returned in the UI
- Includes:
 - Building year
 - Floor size
 - Parcel size
 - Municipality name
 - Neighbourhood name
 - Number of schools within a 3km radius
 - Distance to the nearest train station



ITU KALEIDOSCOPE
ACCRA2022

Different users for data analysis

- GIS developer: GIS tools & GIS formats
- Developer: APIs (only “known” questions)
- (Web) Data Scientists: SPARQL
- Business Analysts: ???
- Citizens: ???

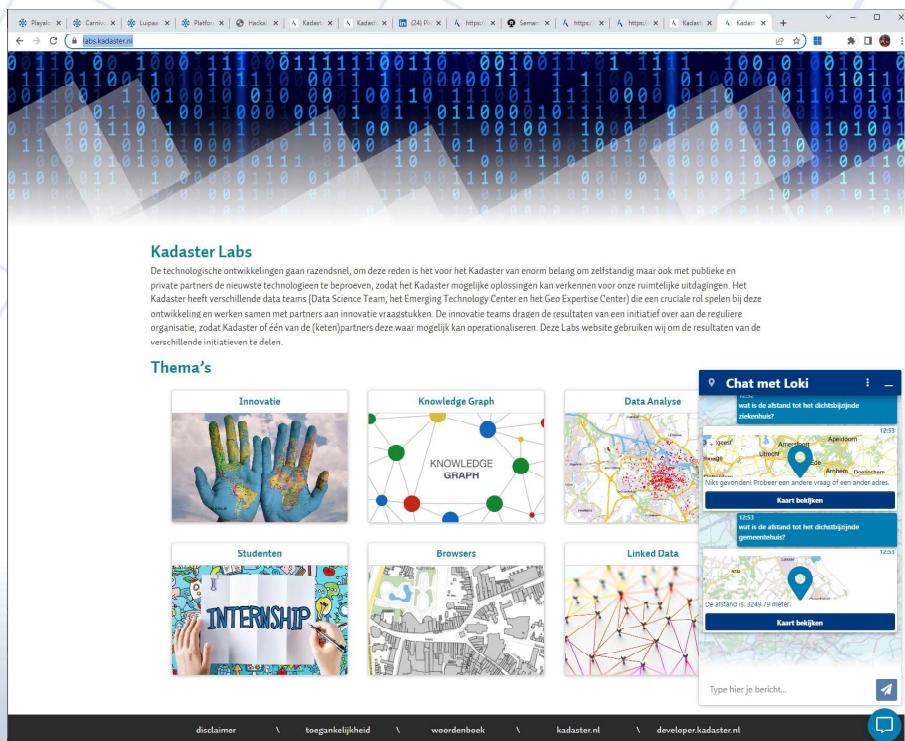
Kadaster Query Builder

The screenshot shows the Kadaster Query Builder interface. At the top, there is a navigation bar with multiple tabs, including 'Home' and 'Lang. v'. Below the navigation bar, the title 'QUERYBUILDER KKG-LIGHT' is displayed. A search bar labeled 'example 1' is present. The main area contains a query graph with three horizontal paths:

- Top path: Gebouw → gebruiksdool → gebruiksdool → gebruiksfunctie.
- Middle path: Gebouw → bouwjaar → bouwjaar → bouwjaar.
- Bottom path: Gebouw → Ligt in → Gemeente → Sinterhogenbosch.

Each path has an 'And' node between the first and second steps. Below the graph is a large blue button with a play icon. At the bottom of the interface, there are tabs for 'Table', 'Chart', 'Geo', 'Pivot', 'Response', and 'Layers'. The 'Geo' tab is selected, displaying a map of a rural area with various locations labeled: Hedel, Crijnsecoeur, Oud Empel, Engelen, Dierck, Empel, Rosmalen, Sprokkelbosch, Heesind, Kruisdonk, Nuland, Geffen, and Lagikant. The map shows roads, water bodies, and land parcels. A legend at the top of the map provides options for 'Normal', 'Grouped', 'Heatmap', and 'Layers'.

But, what if SPARQL is not needed?



Latest Developments

Formal product

Lock Unlock project: Authorization

- Combination of open data
- Changing society
- Adding proprietary data

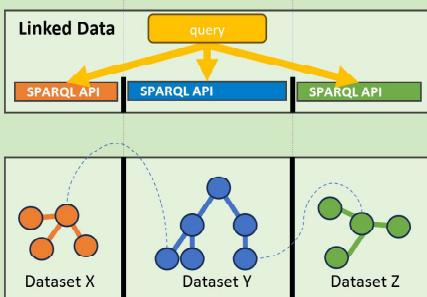
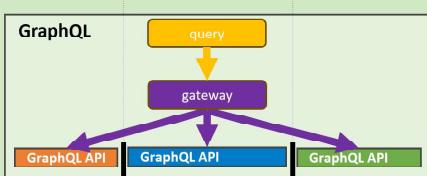
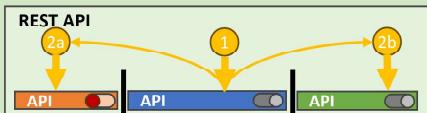
Knowledge Graphs & Large Language Models

Lock-Unlock :lock de data, unlock het potentieel

Versie 0.9 – februari 2024

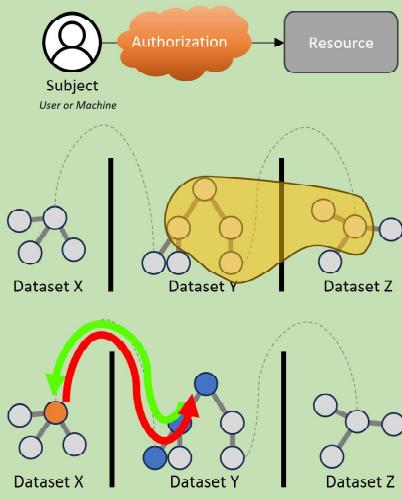


Federatieve bevraging



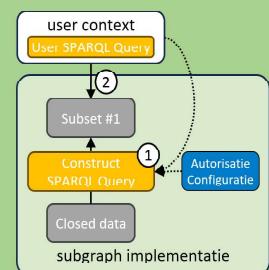
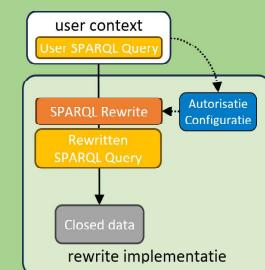
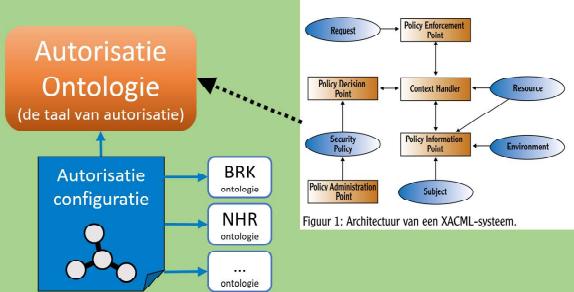
REALISATIE
IBDS FDS Lock-Unlock Project

Afscherming ('lock')



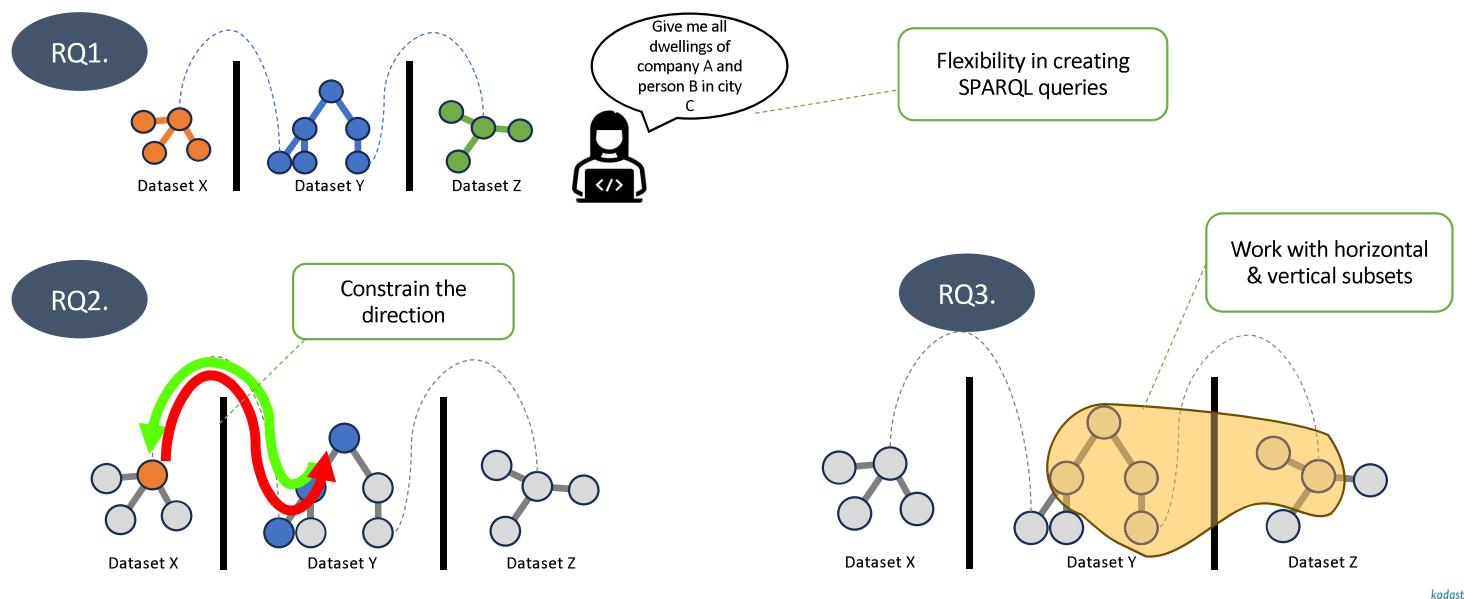
Horizontale subset	?
Verticale subset	?
Richting beperken	?
Vrije query	?

Autorisatie als Linked Data ('unlock')



kadaster

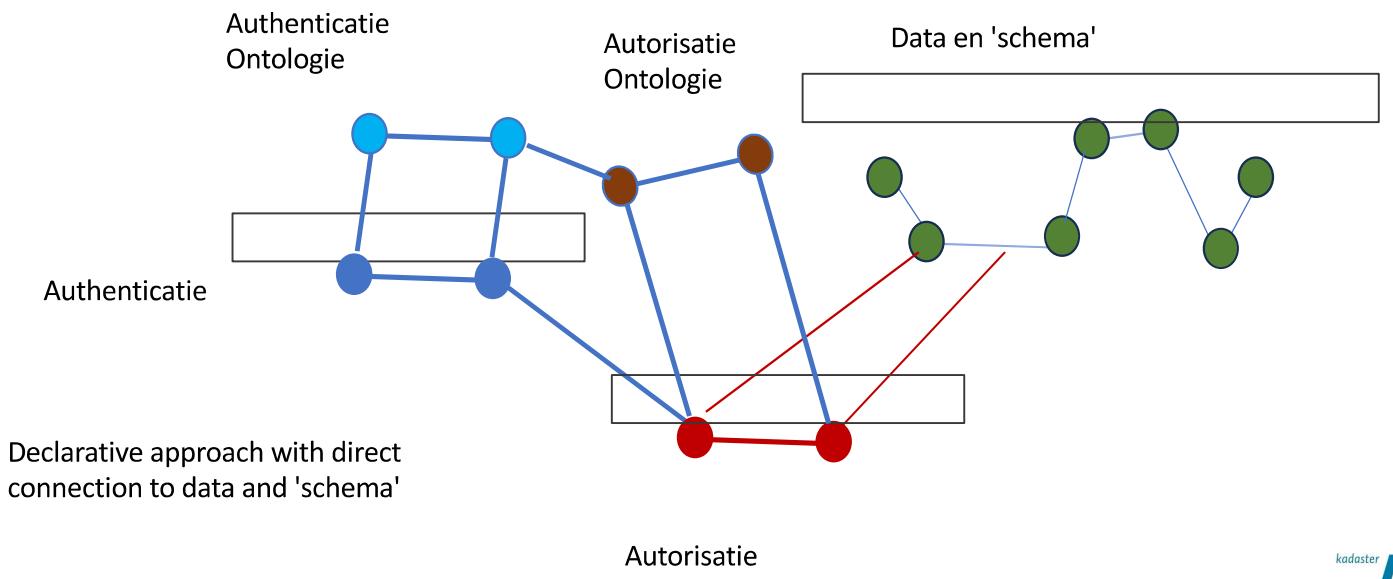
(Advanced) Access Control with Knowledge Graphs



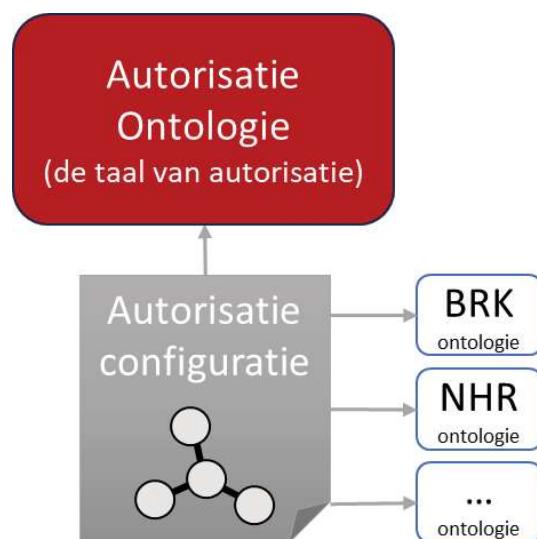
(Advanced) Authorisation in LD

Authorisation ontology

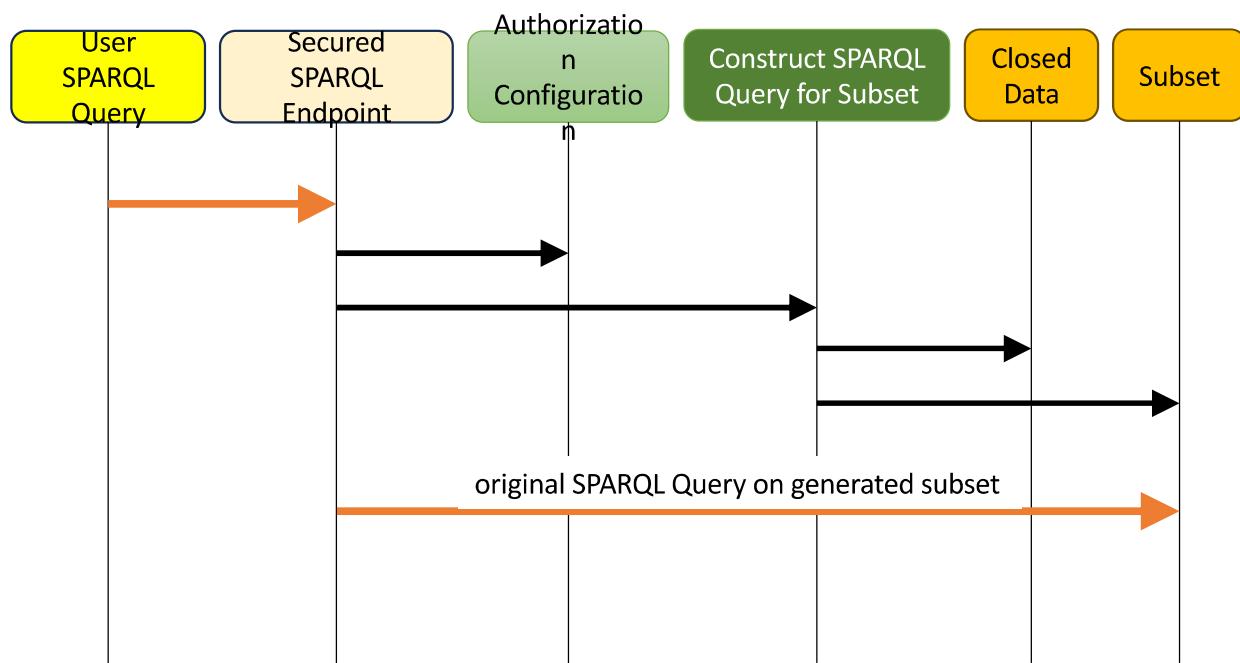
Goal: full Sparql functionality, but no access to closed data



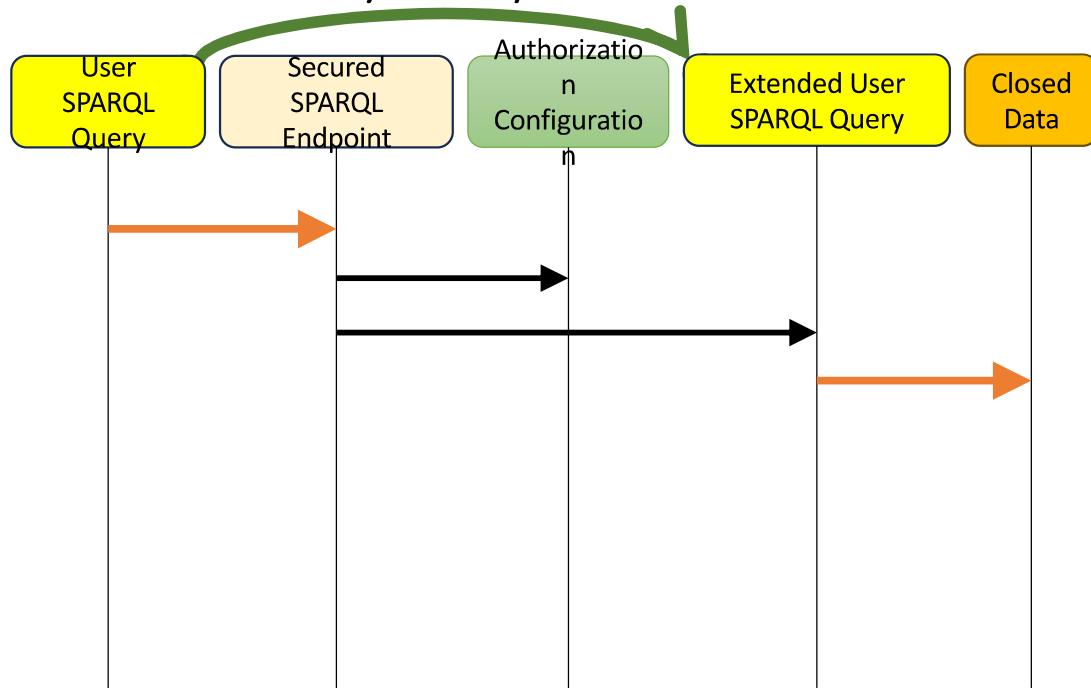
Autorisatie als Linked Data



SPARQL Query story - Subset



SPARQL Query story - Rewrite



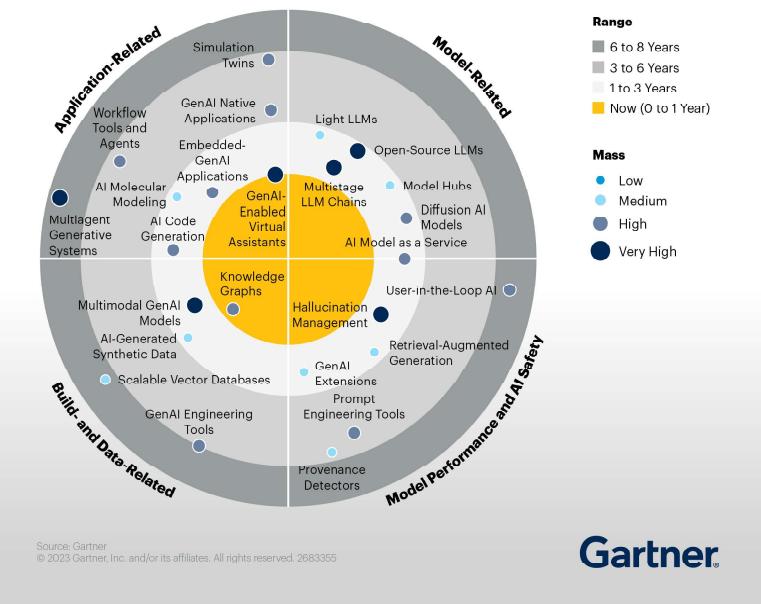
Latest Developments

Lock Unlock project: Authorization

- Combination of open data
- Changing society
- Adding proprietary data

Knowledge Graphs & Large Language Models

Impact Radar for Generative AI



Knowledge Graphs (KGs)

Cons:

- Implicit Knowledge
- Hallucination
- Indecisiveness
- Black-box
- Lacking Domain-specific/New Knowledge

Pros:

- Structural Knowledge
- Accuracy
- Decisiveness
- Interpretability
- Domain-specific Knowledge
- Evolving Knowledge

Large Language Models (LLMs)

Pros:

- General Knowledge
- Language Processing
- Generalizability

Cons:

- Incompleteness
- Lacking Language Understanding
- Unseen Facts



	LLM	KG	KG + LLM
Hallucinations	👎	👍	?
Opaqueness:	👎	👍	?
Intransigence:	👎	👍	?
Staleness:	👎	✊	?
Bias:	👎	✊	?
Costs:	👎	✊	?
Short tail:	👎	👍	?
Sanitised:	👎	👍	?
Non-determinism:	👎	👍	?
Usability:	👍	👎	?

<https://aidanhogan.com/talks/2023-09-22-semantics-2023.pdf>

Why is this important? — Energy consumption



Andreas Blumauer
CEO and Co-Founder

400 metres



~5 seconds



0.56 seconds



0.26 seconds

1 meter

Denny Vrandečić: The Future of Knowledge Graphs
in a World of Large Language Models
<https://www.youtube.com/watch?v=WqYBx2gB6vA>

© Semantic Web Company 2024

OF APPLIED SCIENCES

EN BETROUWBAARHEID VAN HET ANTWOORD?

Wat is het bouwjaar van het huis?

Wat is het bouwjaar van het huis op 't Spiker 24 in Wamsveld?

Copilot

Het bouwjaar van het huis op 't Spiker 24 in Wamsveld is 1976. Dit adres bevindt zich in de buurt Dorp Wamsveld en het huis heeft een energielabel B. De laatste bekende WOZ-waarde van de woning is 571.000 euro. Mocht u meer informatie nodig hebben, sta ik voor u klaar om te helpen!

Meer informatie

1 [huispedia.nl](#) 2 [kadastralekaart.com](#) +3 meer

69

CoPilot
Google: de links...
KG+ Loki: antwoord + links naar bron

Chat met Loki

23:18 wat is het bouwjaar van mijn huis?

23:18 Wat is het adres waar u in geinteresseerd bent?

23:18 't Spiker 24 Wamsveld

Dit pand is gebouwd in 1976

Kaart bekijken

status gebruiksdool perceeloppervlakte ander adres

Type hier je vraag...

SWJ Paper - Google Docs | To Do List - Google Docs | Installatie Erwin Folmer

han.nl/agenda/2024/06/installatie-erwin-folmer/

HAN_

Opleidingen ▾ Studiekeuze ▾ Samenwerken ▾ Onderzoek ▾ NL ▾

Lectoraat Applied Data Science & AI Lectoraten

→ Meld u nu aan!



Op donderdag 21 november wordt Erwin Folmer geïnstalleerd als lector Applied Data Science en AI tijdens het Data & AI event van het Oosten.