



ARCHIXL-PLDN SEMINAR

THESAURI BEST PRACTICES

**PIETER VAN EVERDINGEN
(PLDN/OPENINC)**

30-9-2020

SAMENWERKEN AAN DE NIEUWE PLDN LINKED DATA THESAURUS

PLDN Thesaurus Werkgroep - Update Activiteiten



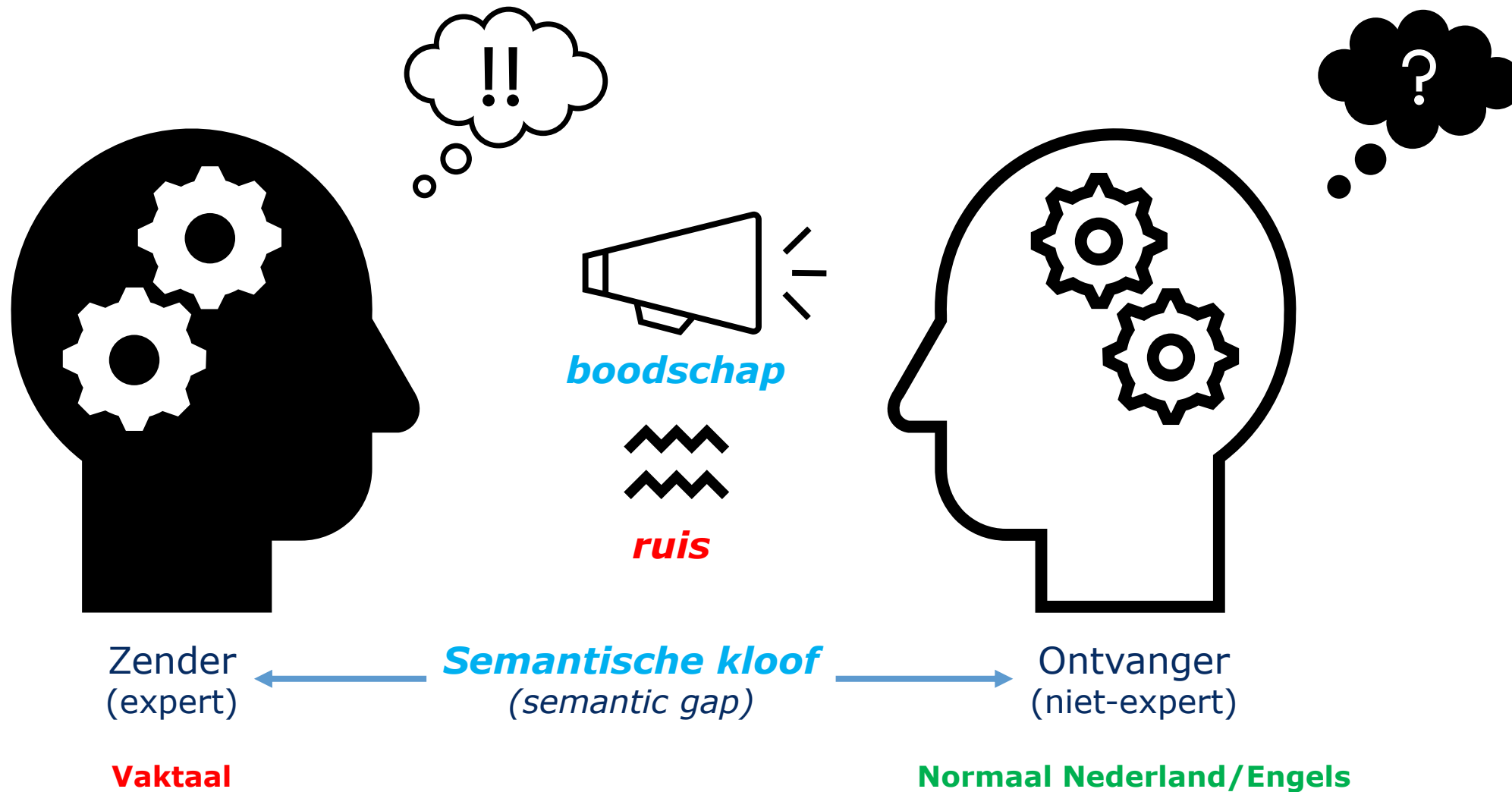
Platform Linked
Data Nederland

Onderwerpen

- Communicatiekloof
- Ambities & doelen
- Aanpak & ontwerpkeuzes
- Workflow & tooling
- Voorlopige resultaat
- Vervolgstappen



Communicatie – Begrijpen we elkaar?





Platform Linked
Data Nederland

PLDN Thesaurus - Werkgroep

Ambitie:

- Linked Data Thesaurus
- Thesaurus Best Practice



30-90 regel*

Doelen:

1. Resultaat (thesaurus)
2. Kennisdeling (ontwerpkeuzen)
3. Gebruiker (aansluiten op behoeften)
4. Onderhoud (workflow & tooling)

**Zo laagdrempelig
mogelijk**

***: Alleen die 30% beschrijven die we 90% van de tijd gebruiken**



PLDN Thesaurus - Maak het vindbaar !!

Voor elk concept (begrip):

- ***Wat is het*** (korte, duidelijke en eenduidige definitie)
- ***Waar kan ik aanvullende informatie vinden*** (W3C, etc.)

Browser/navigatie/zoek-faciliteiten:

1. Zoekveld (zoeken over meerdere thesauri)
2. Alfabetische lijst (zonder verdere structuur)
3. Hiërarchische structuur (fijnmazige classificatie)
4. Thematische structuur (grofmazige groepering)



Bestaande Linked Data thesauri?

Weinig voorbeelden:

- Linked Data Glossary, W3C, 2013 (verouderd)
- Social Semantic Web Thesaurus, 2012 (verouderd)
- GO FAIR Glossary, 2020 (goed als glossary, maar kort)
- ...

Wèl begrippen en definities (als linked data) in:

- DBpedia (137)
- Wikidata (155)
- ...



Platform Linked
Data Nederland

Andere thesauri & ontologieën?

- BBC Ontologies
- Play-a-LOD (model en instanties)
- Getty Vocabularies
- UNESCO Thesaurus (en Publication Ontology)
- AGROVOC
- EuroVoc
- GEMET
- ...



Andere thesauri & ontologieën?

- **BBC Ontologies** (*modulaire opzet*, geen overtollige definities)
- **Play-a-LOD** (schone ontologie, ook geen overtollige definities)
- **Getty Vocabularies** (goed gedocumenteerd, incl. *editorial guidelines*)
- **UNESCO Thesaurus** (*Publication Ontology* t.b.v. thematische structuur)
- **AGROVOC** (gebruiken ook *VocBench*, *Skosmos* en *topconcepten*)
- EuroVoc
- GEMET
- ...



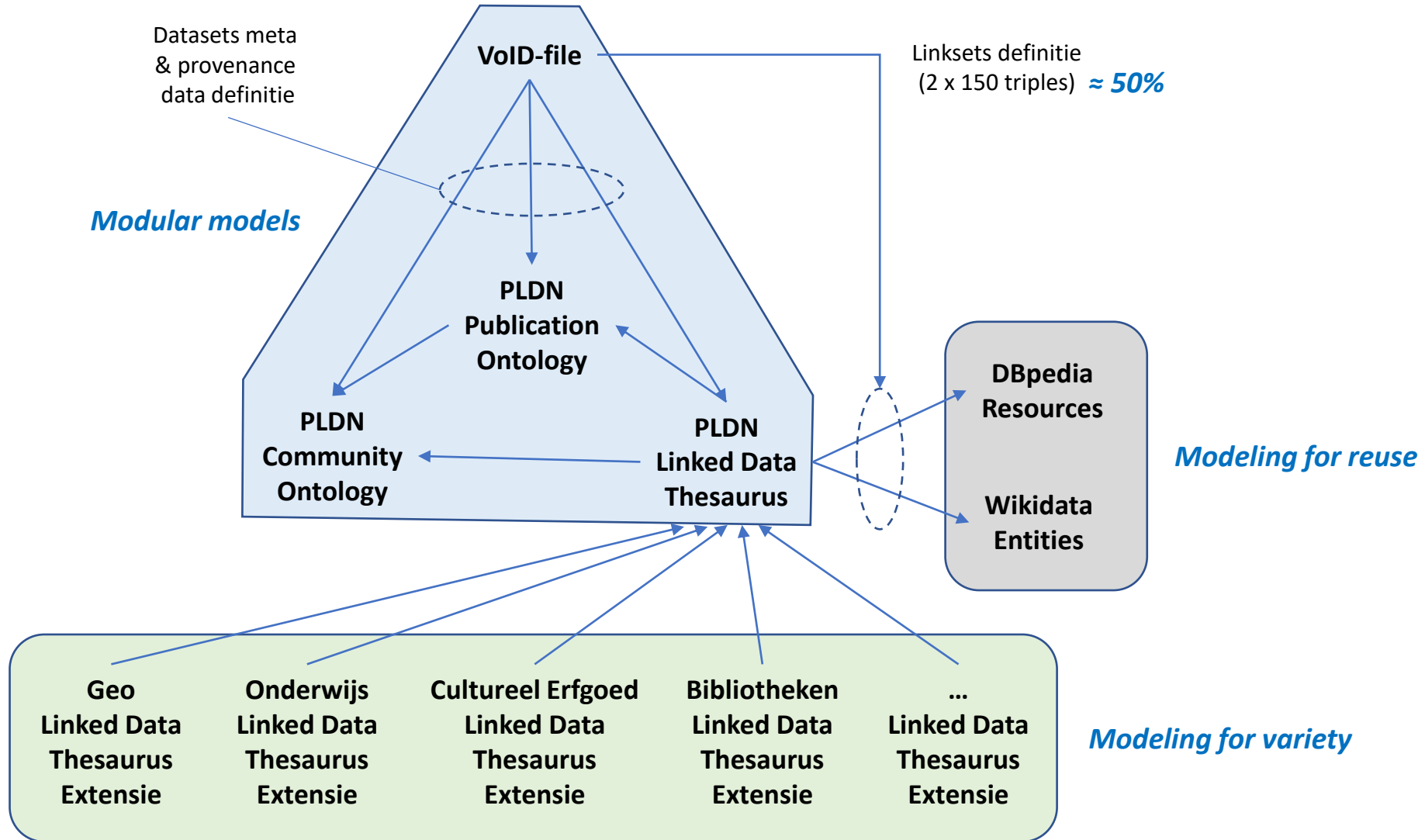
Bestaande Thesauri Best Practices?

Passage uit: ***Semantic Web for the Working Ontologist*** (3rd edition):

Common practice for using these properties is varied: some vocabularies don't use them at all (leaving membership in the concept scheme implicit). Others use *inScheme*, but make no use of top concepts. Among those that indicate to concepts, some use *hasTopConcept* and others use *topConceptOf*.

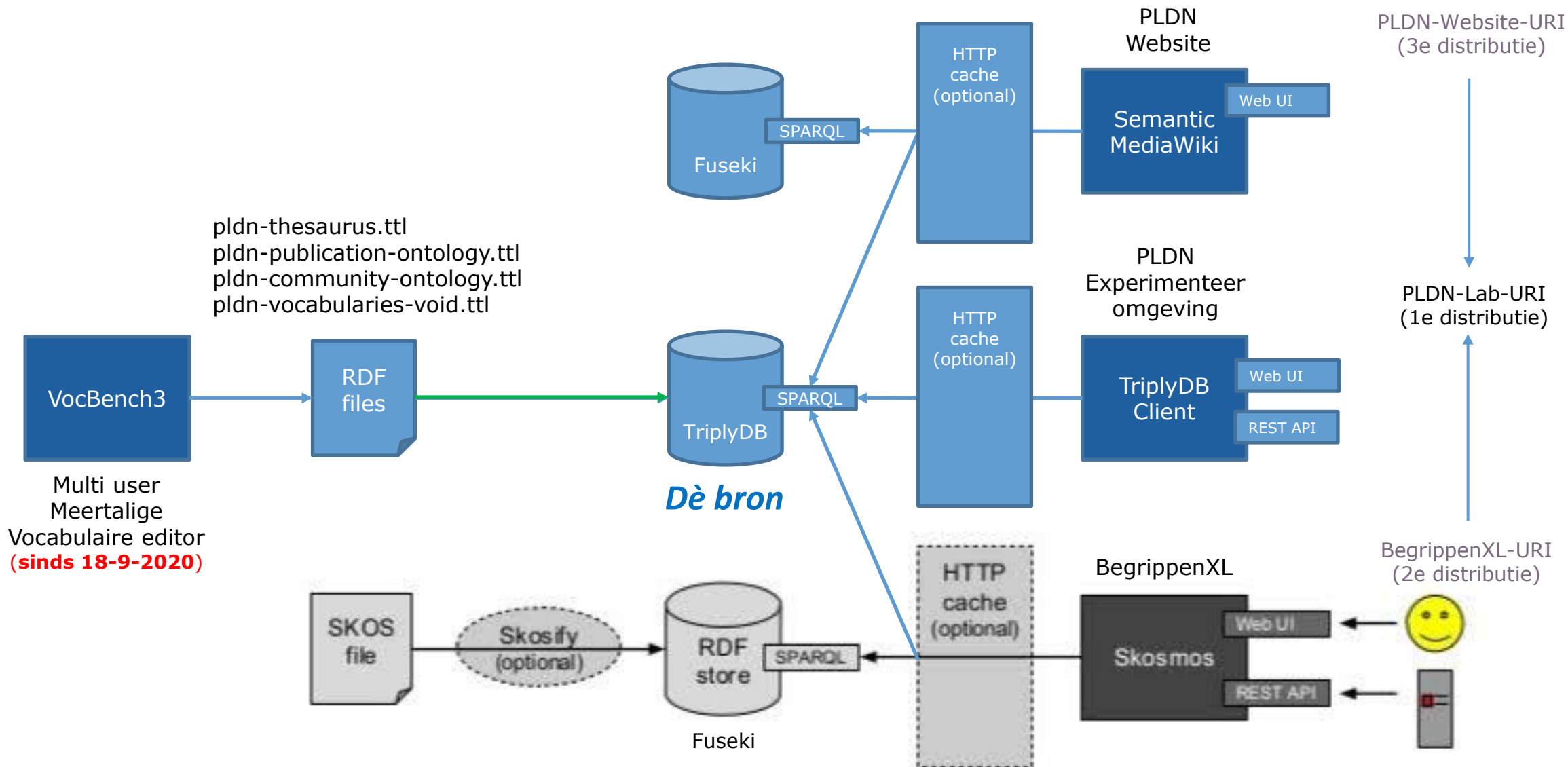
All of these practices are acceptable according to the SKOS standard, but having so many acceptable practices, while making the job of a thesaurus writer easy, makes it more difficult for the consumer of a vocabulary to find their way around.

PLDN Thesaurus – Niet stand-alone





PLDN Thesaurus - Gewenste workflow





PLDN Thesaurus – Concepts template

```
File Edit Selection View Go Run Terminal Help pldn-thesaurus-concept-template.ttl - Visual Studio Code
unesco.ttl pldn-thesaurus-concept-template.ttl x pldn-thesaurus-review-version.ttl pldn-thesaurus.ttl pldn-publication-o

1 <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/...>
2 a skos:Concept ;
3 dcterms:modified "2020-09-19"^^xsd:date ;
4 skos:prefLabel "..."@en , "..."@nl ;
5 skos:altLabel "..."@en , "..."@nl ;
6 vaem:abbreviation "..."@en , "..."@nl ;
7 skos:definition "..."@en ;
8 skos:definition "..."@nl ;
9 skos:inScheme <https://data.pldn.nl/pldn-vocabularies/Thesaurus> ;
10 skos:topConceptOf <https://data.pldn.nl/pldn-vocabularies/Thesaurus> ;
11 isother:broaderGeneric <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;
12 isother:broaderPartitive <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;
13 isother:broaderInstantial <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;
14 skos:related <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;
15 skos:exactMatch <http://dbpedia.org/resource/...> ;
16 skos:closeMatch <http://dbpedia.org/resource/...> ;
17 skos:exactMatch <https://www.wikidata.org/entity/...> ;
18 skos:closeMatch <https://www.wikidata.org/entity/...> ;
19 foaf:primaryTopicOf <https://...> ;
20 dcterms:source "..."@en , "..."@nl ;
21 .

Ln 21, Col 2 Spaces: 2 UTF-8 CRLF TURTLE 09:49 30-9-2020
```



PLDN Thesaurus – Concept voorbeeld

```
File Edit Selection View Go Run Terminal Help • pldn-thesaurus-review-version.ttl - Visual Studio Code
unesco.ttl pldn-thesaurus-concept-template.ttl pldn-thesaurus-review-version.ttl • pldn-thesaurus.ttl pldn-publication-ontology.ttl pldn-community ...
1020
1021 <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/SimpleKnowledgeOrganizationSystem>
1022   a skos:Concept ;
1023   dcterms:modified "2020-09-28"^^xsd:date ;
1024   skos:prefLabel "SKOS"@en , "SKOS"@nl ;
1025   skos:altLabel "Simple Knowledge Organization System"@en ;
1026   skos:altLabel "Simple Knowledge Organization System"@nl ;
1027   vaem:abbreviation "SKOS"@en , "SKOS"@nl ;
1028   skos:definition "RDF-based vocabulary that can be used to define taxonomies and thesauri"@en ;
1029   skos:definition "RDF-gebaseerde vocabulaire die gebruikt kan worden om taxonomieën en thesauri mee te definiëren"@nl ;
1030   isother:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/RDF-BasedVocabularies> ;
1031   isother:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/OpenStandards> ;
1032   isother:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/W3C-Recommendations> ;
1033   isother:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/FS-MandatoryStandards> ;
1034   skos:related <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/Taxonomies> ;
1035   skos:related <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/Thesauri> ;
1036   skos:related <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/InformalModels> ;
1037   skos:inScheme <https://data.pldn.nl/pldn/vocabularies/Thesaurus> ;
1038   skos:exactMatch <http://dbpedia.org/resource/Simple_Knowledge_Organization_System> ;
1039   skos:exactMatch <https://www.wikidata.org/entity/Q2288360> ;
1040   foaf:primaryTopicOf <https://www.w3.org/TR/skos-primer/> ;
1041   foaf:primaryTopicOf <https://www.w3.org/TR/skos-reference/> ;
1042   foaf:primaryTopicOf <https://www.forumstandaardisatie.nl/open-standaarden/skos> ;
1043   dcterms:source "W3C"@en , "Forum Standaardisatie"@en , "DBpedia"@en , "Wikidata"@en ;
1044   dcterms:source "W3C"@nl , "Forum Standaardisatie"@nl , "DBpedia"@nl , "Wikidata"@nl ;
1045
1046
1047
1048
1049
```

Ln 1045, Col 2 Spaces: 2 UTF-8 CRLF TURTLE 99% 14:52 1-10-2020



PLDN Thesaurus – In Vocbench3 (PLDN)

The screenshot shows the VocBench3 interface for the PLDN Thesaurus project. The top navigation bar includes 'About VocBench', 'Current project: PLDN_Thesaurus', and 'Global Data Management'. The main menu has 'VocBench', 'Projects', 'Data', 'Metadata', 'SPARQL', and 'Tools'. The 'Scheme' tab is selected and highlighted with a red box. The left pane shows a list of schemes with 'PLDN Thesaurus (en), PLDN Thesaurus (nl)' selected. The right pane displays the configuration for the selected scheme, including 'Types' (rdf:type: skos:ConceptScheme), 'Lexicalizations' (skos:prefLabel: PLDN Thesaurus), and 'Notes'. A large blue text overlay 'Nu 1 scheme' is positioned over the 'Types' section. The bottom status bar shows 'Status: K' and navigation buttons 'ResView' and 'Code'.



PLDN Thesaurus – In Vocbench3 (PLDN)

The screenshot displays the VocBench3 interface for the PLDN Thesaurus. The left sidebar shows a tree view of concepts, with 'Data on The Web Best Practices (en)' selected. The main panel shows the configuration for this concept, including its type (skos:Concept), scheme (PLDN Thesaurus), and broader concept (Best Practices). The text 'Nu 468 concepten en 22 topconcepten' is overlaid on the main panel.

**Nu 468 concepten
en 22 topconcepten**



PLDN Thesaurus – Topconcepten (22)

- > Applications
- > Assumptions
- > Best Practices
- > Data
- > Formats
- > Identifiers
- > Interfaces
- > Interoperability
- > Languages
- > Mappings
- > Models
- > Principles
- > Processes
- > Profiles
- > Protocols
- > RDF Ecosystem
- > Reasoning
- > Standards
- > Strategies
- > Tooling
- > Vocabularies
- > World Wide Web

PLDN thesaurus

Content language **English** Search

- Alphabetical **Hierarchy** Groups
- > Applications
 - > Assumptions
 - > Best Practices
 - > Data
 - > Formats
 - > Identifiers
 - > Interfaces
 - > Interoperability
 - > Languages
 - > Mappings
 - > Models
 - > Principles
 - > Processes
 - > Profiles
 - > Protocols
 - > RDF Ecosystem
 - > Reasoning
 - > Standards
 - > Strategies
 - > Tooling
 - > Vocabularies
 - > World Wide Web

Mock-up

Vocabulary information

PREFERRED TERM	pldn thesaurus
TYPE	http://www.w3.org/2004/02/skos/core#ConceptScheme
URI	https://www.begrippenxl.nl/pldn/nl/page/

Resource counts by type

Type	Count
Concept	468

Term counts by language

Language	Preferred terms	Alternate terms	Hidden terms
English	468	1	0
Dutch	468	1	0

PLDN thesaurus

Content language **Dutch** Search

- Alphabetical **Hierarchy** Groups
- > Aannames
 - > Best Practices
 - > Data
 - > Formaten
 - > Identifiers
 - > Interfaces
 - > Interoperabiliteit
 - > Mappingen
 - > Modellen
 - > Principes
 - > Processen
 - > Profielen
 - > Protocollen
 - > RDF Ecosysteem
 - > Redeneren
 - > Standaarden
 - > Strategieën
 - > Talen
 - > Toepassingen
 - > Tooling
 - > Vocabulaires
 - > World Wide Web

Mock-up

Vocabulary information

TITLE	PLDN thesaurus
TYPE	http://www.w3.org/2004/02/skos/core#ConceptScheme
URI	https://www.begrippenxl.nl/pldn/nl/page/

Resource counts by type

Type	Count
Concept	468

Term counts by language

Language	Preferred terms	Alternate terms	Hidden terms
English	468	1	0
Dutch	468	1	0

Volledig tweetalig



PLDN Thesaurus – Stand van zaken (30/9)

We hebben alle puzzlestukjes, maar we moeten nu alleen de puzzle nog 'even' in elkaar zetten





Platform Linked
Data Nederland

PLDN Thesaurus – Hoe nu verder?

PLDN Thesaurus Werkgroep

- Pieter van Everdingen
- Hans de Rond
- Pano Maria
- Jan Willem van Veen

Nieuwe deelnemers aanmelden via:

- info@pldn.nl



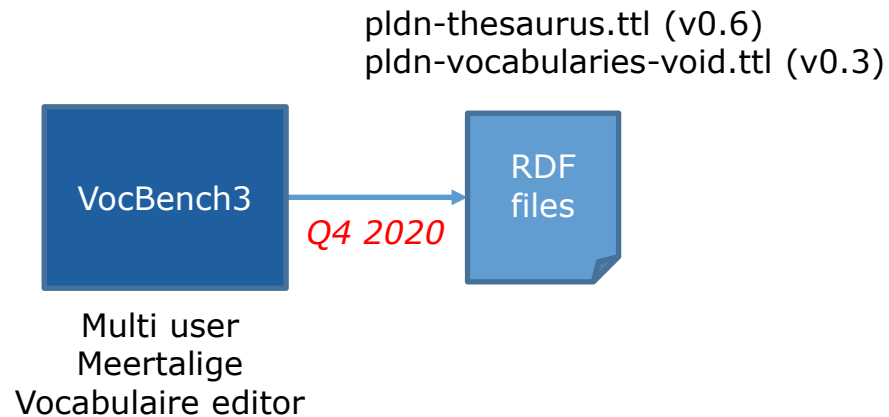
PLDN Thesaurus – Hoe nu verder?

PLDN Thesaurus Werkgroep

- Editors toevoegen
- Reviewers toevoegen

Eerste activiteiten:

- Werkproces afspreken
- Documentatieproces afspreken
- Naamgevingsconventie afspreken
- URI-strategie afspreken
- Eerste thesaurusopzet reviewen
- Feedback verwerken in volgende versie van de thesaurus





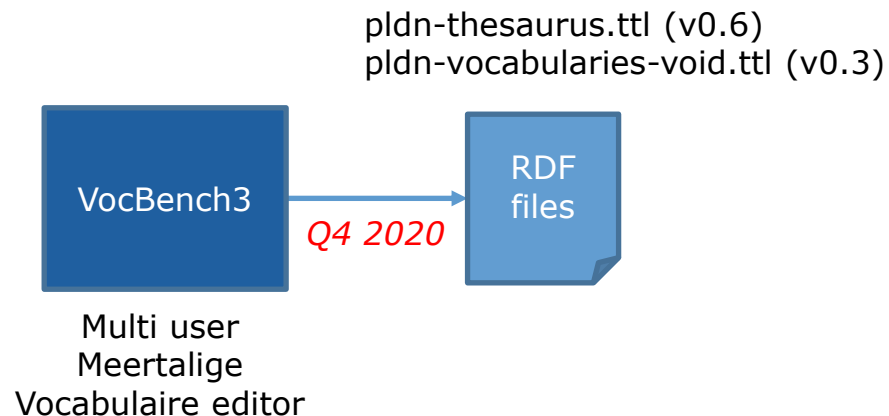
PLDN Thesaurus – Hoe nu verder?

PLDN Thesaurus Werkgroep

- Editors toevoegen
- **Reviewers toevoegen**

Eerste activiteiten:

- **Werkproces afspreken**
- Documentatieproces afspreken
- Naamgevingsconventie afspreken
- URI-strategie afspreken
- **Eerste thesaurusopzet reviewen**
- Feedback verwerken in volgende versie van de thesaurus





Platform Linked
Data Nederland

Werkproces – Reqs & issues op Github?

The screenshot shows the GitHub interface for the organization 'Platform Linked Data Nederland'. At the top, there is a navigation bar with the GitHub logo, a search bar, and links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. Below the navigation bar, the organization's profile is displayed, including the organization's logo and name. A horizontal menu below the profile contains tabs for 'Repositories', 'Packages', 'People', and 'Projects'. The 'Repositories' tab is currently selected. The main content area shows a message: 'This organization has no public repositories.' To the right of this message, there is a 'People' section with a right-pointing arrow. The 'People' section contains the text: 'This organization has no public members. You must be a member to see who's a part of this organization.' Below the 'People' section, there is a 'Report abuse' link. At the bottom of the page, there is a footer with copyright information and various links: '© 2020 GitHub, Inc.', 'Terms', 'Privacy', 'Security', 'Status', 'Help', 'Contact GitHub', 'Pricing', 'API', 'Training', 'Blog', and 'About'.



Platform Linked
Data Nederland

Werkproces – Ontwerpkeuzen vastleggen



Platform Linked
Data Nederland

Innovatie door betekenisvol verbinden

Wat zoekt u?



platform

linked data

experts

doe mee

PLDN Thesaurus Best Practices

PLDN Werkdocument 27 september 2020

Deze versie:

https://www.pldn.nl/wiki/PLDN_Thesaurus_Best_Practices

Redacteur:

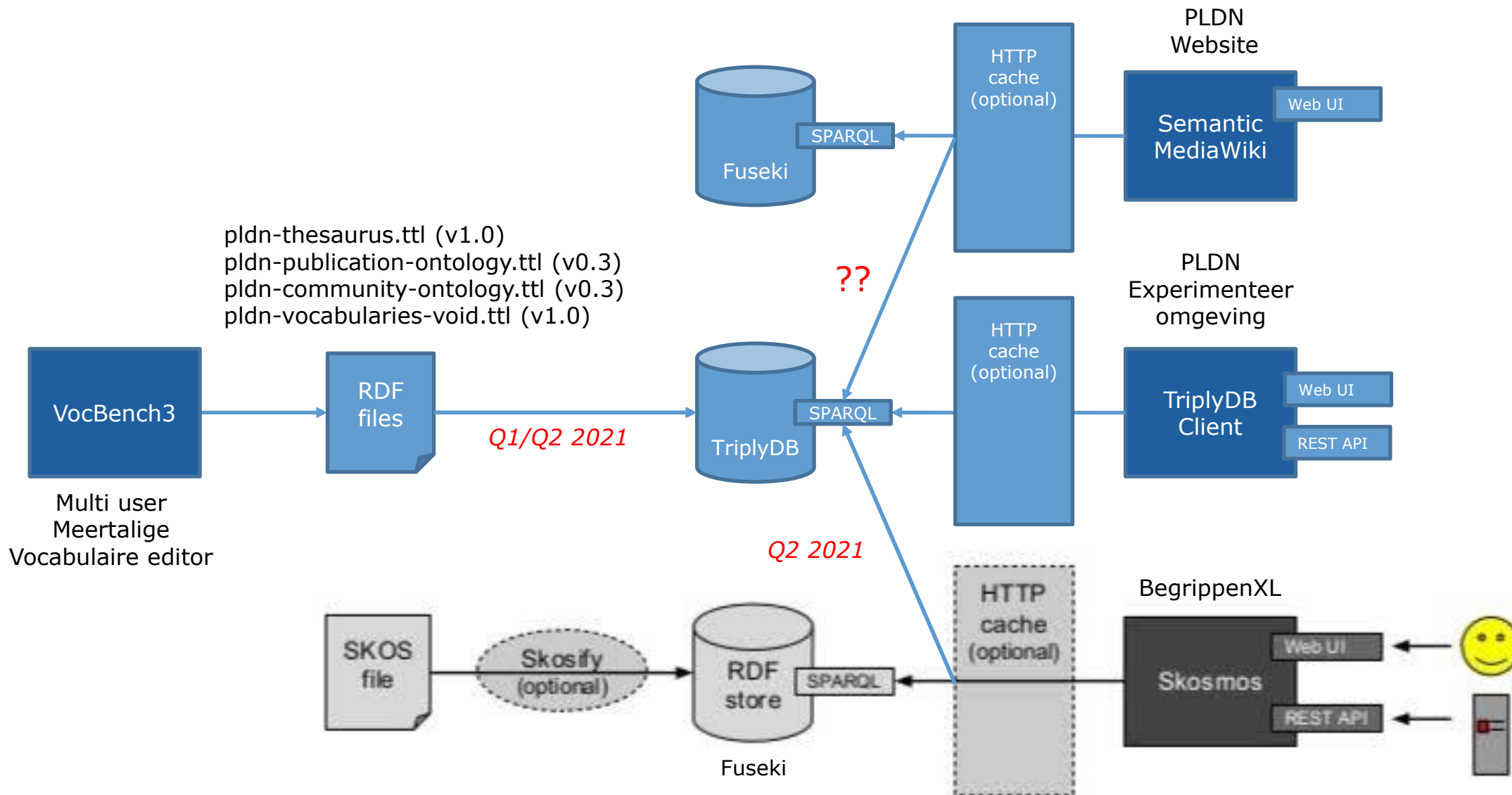
Pieter van Everdingen, PLDN

Reviewers:

Hans de Rond, Bureau De Rond
Pano Maria, Skemu



PLDN Thesaurus – Hoe nu verder?





PLDN Thesaurus – Take aways

- Domeinkennis, domeinkennis, domeinkennis...
- Verzamel zoveel mogelijk begrippen (zonder grens) (dit kan helpen om patronen makkelijker te zien)
- Ga rigoureuus schrappen om tot de kern te komen
- Als een structuur niet werkt, begin dan opnieuw (niet 'doormodderen' als je niet tevreden bent)
- Ga pas een tool gebruiken als je wel tevreden bent
- Luister naar de oude thesaurus rotten in dit vak
- Het is vooral veel handwerk... (menselijke keuzes)





Platform Linked
Data Nederland

PLDN communicatiekanalen

BIG DATA
VALUE CENTER
OPEN INNOVATION CENTER

ECP

Platform voor de
InformatieSamenleving

Contact

Pieter van Everdingen

info@pldn.nl



Website

www.pldn.nl

Experimenteeromgeving

www.data.pldn.nl

LinkedIn-group

www.linkedin.com/groups/466278

Twitter [@linkeddatanl](https://twitter.com/linkeddatanl)

hashtag #LinkedDataNL

Nieuwsbrief

www.pldn.nl/wiki/Nieuwsbrieven



Platform Linked
Data Nederland

BACKUP SLIDES





PLDN Vocabulaires – URI-strategie

- <https://data.pldn.nl/pldn/vocabularies/CommunityOntology>
- <https://data.pldn.nl/pldn/vocabularies/PublicationOntology>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus>

- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/LinkedData>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/RDF>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/SPARQL>
- [...](#)

- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/LinkedDataCore>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Abbreviations>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Standards>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Data>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Tooling>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Guidance>
- <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/group/Licenses>

Elke concept group heeft minimaal enkele tientallen en maximaal enkele honderden termen (huidige inschatting)



Innovatie door betekenisvol verbinden

- platform
 - linked data
 - experts**
 - doe mee
- experts
 - expertise
 - organisaties
 - blogs

Expertises **Oude versie**

Welke expertise is er zoal aanwezig binnen het Platform Linked Data Nederland? Onderstaand overzicht geeft een beeld van de onderwerpen en bijbehorende experts. Deze lijst is nog in bewerking en zal stap-voor-stap nog verder uitgewerkt worden met meer detail.

3

- [303 redirect](#)
- [3D geo-informatie](#)

4

- [4 Linked Data principes](#)

5

D meer

- [Data replicatie](#)
- [Data schema's](#)
- [Data science](#)
- [Data scraping](#)
- [Data services](#)
- [Data sets](#)
- [Data shapes](#)

L meer

- [LIMES](#)
- [LOD Cloud](#)
- [LOD Cloud Knowledge Graph vs COVID-19](#)
- [LOD Laundromat](#)
- [LOD2 Stack](#)
- [LOTUS](#)
- [LOV](#)

PLDN thesaurus

Content language English Search

Alphabetical Hierarchy Groups

A B C D E F G H I J K L
M N O P Q R S T U V W Y
Z 0-9

- A-box reasoning
- Access Control Lists
- Activity Vocabulary
- ADMS
- AI
- AllegroGraph
- Ambiguity
- AnzoGraph
- Apache Jena
- API Gateways
- API Management
- API's
- API-first
- API-strategy
- Application Profiles
- Approaches
- Apps
- AR/VR
- ARC2
- Aspects
- Accepted triples

Vocabulary information

Oude versie

PREFERRED TERM pldn thesaurus

TYPE <http://www.w3.org/2004/02/skos/core#ConceptScheme>

URI <https://www.begrippenxl.nl/pldn/nl/page/>

Resource counts by type

Type	Count
Concept	1159

Term counts by language

Language	Preferred terms	Alternate terms	Hidden terms
English	408	1	0
Dutch	1159	1	0