



WORKING TOGETHER ON THE NEW PLDN LINKED DATA THESAURUS

PLDN Thesaurus Working Group - Update Activities



ARCHIXL-PLDN WEBINAR THESAURI BEST PRACTICES

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

30-9-2020



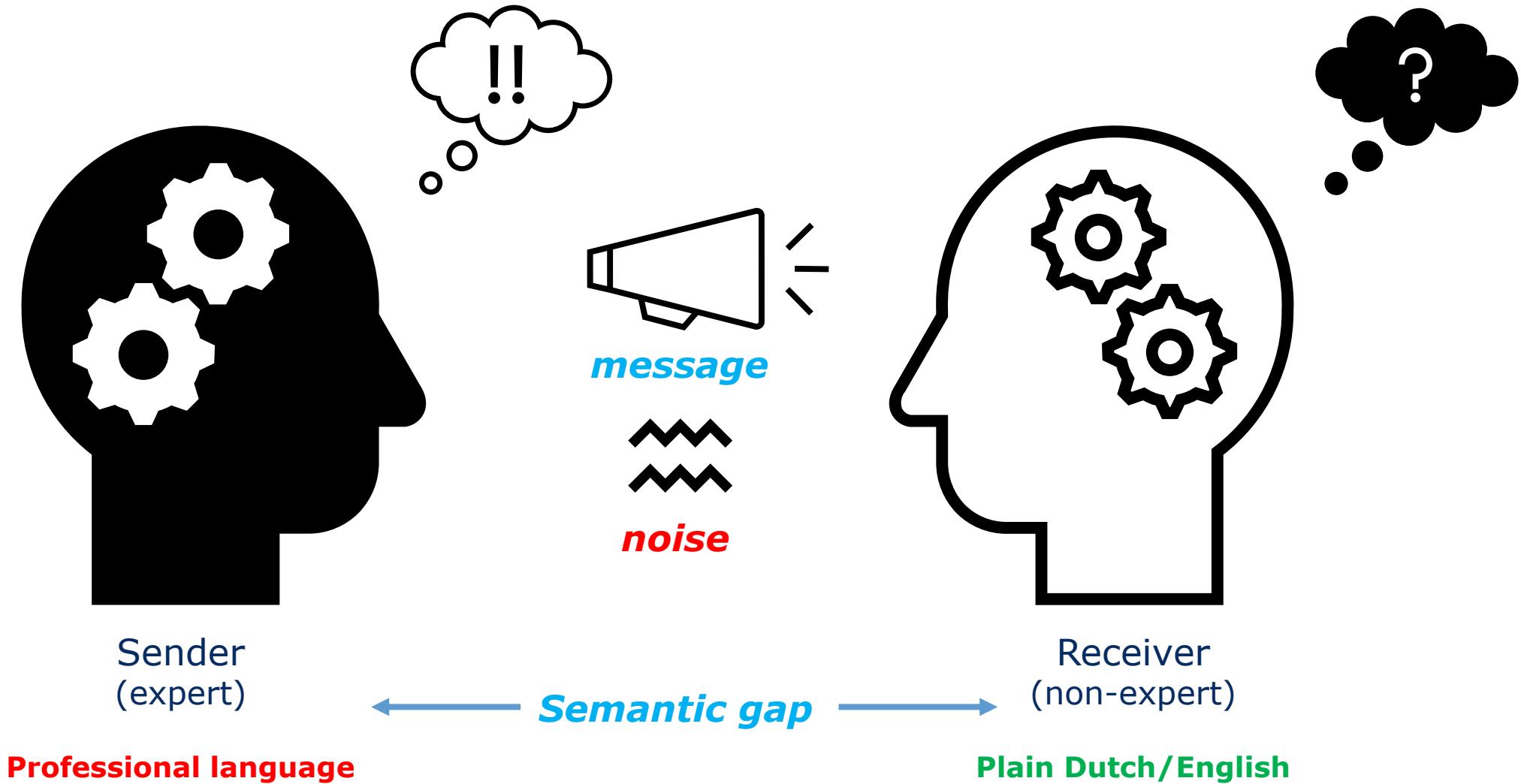
Platform Linked
Data Nederland

Topics

- Communication gap
- Ambitions & goals
- Approach & design decisions
- Workflow & tooling
- Intermediate result
- Next steps



Communication – Do we understand each other??





Platform Linked
Data Nederland

Reduce the communication gap via e.g.

Gratis verzending vanaf 20,- Bezorging dezelfde dag, 's avonds of in het weekend* Gratis retourneren Select Ontdek het nu voor 9.99 p.j.

bol.com Waar ben je naar op zoek? Alles Inloggen

Categorieën ▾ Cadeaus & Inspiratie ▾ Aanbiedingen ▾ Verkopen Zakelijk Cadeaukaart Bestelstatus Klantenservice NL ▾

Alles voor op reis Buitenspeelgoed Fashion pre-sale **Afkoelen** tijdens de hitte! >

Play-a-LOD Dieren

Schrijf een review | E-mail deze pagina



15:

5 - 7 dagen ⓘ Verkoop door partner van bol.com PLDN ⓘ

+ In winkelwagen Op verlanglijstje

✓ Gratis verzending
✓ 30 dagen bedenkijd en gratis retourneren

Vragen en antwoorden (0)

<  >

Learn playfully the basic concepts of linked data via an animal ontology

<https://www.playalod.nl>



PLDN Linked Data Thesaurus Working Group

Ambition:

- Linked Data Thesaurus
- Thesaurus Best Practice



30-90 rule*

Goals:

As easy as possible

1. Result (thesaurus)
2. Knowledge sharing (design decisions)
3. User (user-centric, fit for use)
4. Maintenance (workflow & tooling)

***: Only describe the 30% that we use 90% of the time**



PLDN Thesaurus – Make it findable !!

For each concept (notion):

- ***What does it mean*** (short, clear and concise definition)
- ***Where can I find additional information*** (W3C, etc.)

Browser/navigation/search-facilities*:

1. Search field (search across multiple thesauri)
2. Alphabetic list (without further structure)
3. Hierarchical structure (fine-grained classification)
4. Thematic structure (coarse-grained grouping)

***: Show all possible design decisions with a small thesaurus**



Existing Linked Data thesauri?

Few examples :

- Linked Data Glossary, W3C, 2013 (deprecated)
- Social Semantic Web Thesaurus, 2012 (deprecated)
- GO FAIR Glossary, 2020 (ok as glossary, but small)
- ...

But concepts and definitions (as linked data) in:

- DBpedia (137)
 - Wikidata (155)
 - ...
- } of our 308 core linked data concepts



Other thesauri & ontologies?

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



Other thesauri & ontologies?

- BBC Ontologies (*modular structure*, no redundant definitions)
- Play-a-LOD (clean ontology, also no redundant definitions)
- Getty Vocabularies (well documented, incl. *editorial guidelines*)
- UNESCO Thesaurus (*Publication Ontology* for thematic structure)
- AGROVOC (also use *VocBench*, *Skosmos* and *top concepts*)
- EuroVoc
- GEMET
- ...



Existing Thesauri Best Practices?

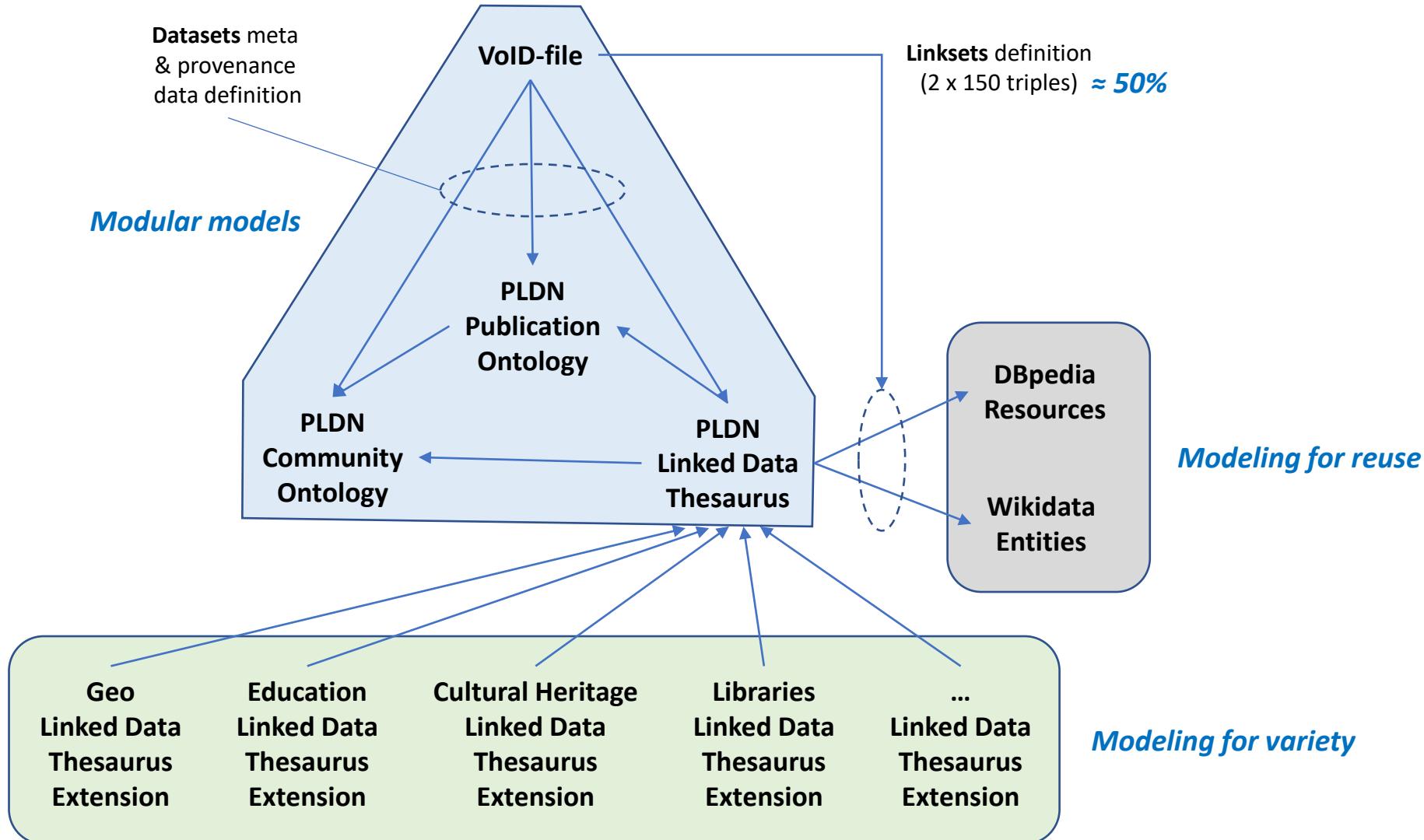
Source: ***Semantic Web for the Working Ontologist*** (3rd edition):

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.



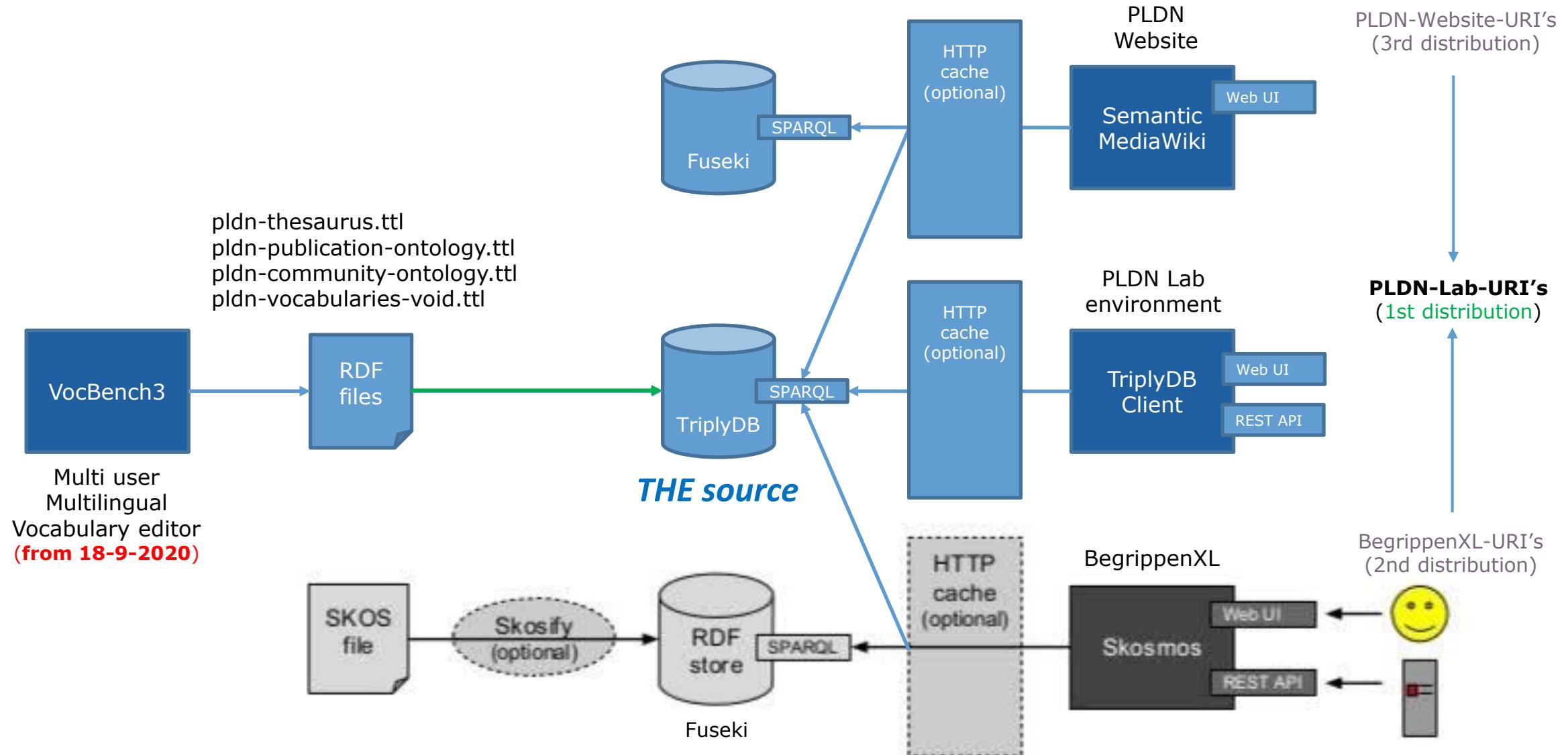
PLDN Thesaurus – Not stand-alone





Platform Linked
Data Nederland

PLDN Thesaurus - Preferred workflow





PLDN Thesaurus – Concepts template

The screenshot shows a Visual Studio Code window with the title "pldn-thesaurus-concept-template.ttl - Visual Studio Code". The left sidebar contains icons for file operations like Open, Save, Find, and Settings. The main editor area displays an RDF triples file. Several sections of the code are highlighted with red boxes:

- Line 6: `vaem:abbreviation "..."@en , ..."@nl ;`
- Line 10: `skos:topConceptOf <https://data.pldn.nl/pldn-vocabularies/Thesaurus> ;`
- Line 11: `isothes:broaderGeneric <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;`
- Line 12: `isothes:broaderPartitive <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;`
- Line 13: `isothes:broaderInstantial <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;`
- Line 14: `skos:related <https://data.pldn.nl/pldn-vocabularies/Thesaurus/id/concept/...> ;`
- Line 15: `skos:exactMatch <http://dbpedia.org/resource/...> ;`
- Line 16: `skos:closeMatch <http://dbpedia.org/resource/...> ;`
- Line 17: `skos:exactMatch <https://www.wikidata.org/entity/...> ;`
- Line 18: `skos:closeMatch <https://www.wikidata.org/entity/...> ;`
- Line 19: `foaf:primaryTopicOf <https://...> ;`
- Line 20: `dcterms:source "..."@en , ..."@nl ;`



PLDN Thesaurus – Concept example

The screenshot shows a Visual Studio Code window with several tabs open. The active tab is "pldn-thesaurus-review-version.ttl". The code displays RDF triples for a "SimpleKnowledgeOrganizationSystem" concept. Several triples are highlighted with red boxes:

- Line 1024: `skos:prefLabel "SKOS"@en , "SKOS"@nl ;`
- Line 1027: `vaem:abbreviation "SKOS"@en , "SKOS"@nl ;`
- Line 1030 to 1033: A block of four triples related to broader instantiations:
 - `isothes:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/RDF-BasedVocabularies> ;`
 - `isothes:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/OpenStandards> ;`
 - `isothes:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/W3C-Recommendations> ;`
 - `isothes:broaderInstantial <https://data.pldn.nl/pldn/vocabularies/Thesaurus/id/concept/FS-MandatoryStandards> ;`
- Line 1038 and 1039: Two triples indicating exact matches:
 - `skos:exactMatch <http://dbpedia.org/resource/Simple_Knowledge_Organization_System> ;`
 - `skos:exactMatch <https://www.wikidata.org/entity/Q2288360> ;`
- A large block of triples from line 1040 to 1044, all enclosed in a single red box:
 - `foaf:primaryTopicOf <https://www.w3.org/TR/skos-primer/> ;`
 - `foaf:primaryTopicOf <https://www.w3.org/TR/skos-reference/> ;`
 - `foaf:primaryTopicOf <https://www.forumstandaardisatie.nl/open-standaarden/skos> ;`
 - `dcterms:source "W3C"@en , "Forum Standaardisatie"@en , "DBpedia"@en , "Wikidata"@en ;`
 - `dcterms:source "W3C"@nl , "Forum Standaardisatie"@nl , "DBpedia"@nl , "Wikidata"@nl ;`

The status bar at the bottom shows: Ln 1045, Col 2 Spaces: 2 UTF-8 CRLF TURTLE. The taskbar at the bottom includes icons for various applications like File Explorer, Task View, and Edge.



Platform Linked
Data Nederland

PLDN Thesaurus – In Vocbench3 (PLDN)

About VocBench ▾ Current project: PLDN_Thesaurus Global Data Management ▾

VocBench Projects Data Metadata ▾ SPARQL Tools ▾

Class Concept **Scheme** Collection Property Datatype

OR A C

PLDN Thesaurus (en), PLDN Thesaurus (nl)

Search... ResView Code

PLDN Thesaurus (en), PLDN Thesaur...

PLDN Thesaurus (en), PLDN Thesaur... https://data.pldn.nl/pldn/vocabularies/Thesaurus/ conceptScheme Rename A C

Now 1 scheme

Types:

rdf:type

skos:ConceptScheme

Lexicalizations:

skos:prefLabel

PLDN Thesaurus

PLDN Thesaurus

Notes:

Status: ✓

The screenshot shows the VocBench3 interface for managing the PLDN Thesaurus. On the left, there's a search bar and a sidebar with tabs for Class, Concept, Scheme (which is highlighted with a red box), Collection, Property, and Datatype. Below these are various icons and a search bar. The main right-hand panel displays a single concept scheme named 'PLDN Thesaurus (en), PLDN Thesaur...' with its URI 'https://data.pldn.nl/pldn/vocabularies/Thesaurus/'. The 'conceptScheme' field is populated. The interface includes sections for Types (containing 'rdf:type' and 'skos:ConceptScheme'), Lexicalizations (containing two entries for 'PLDN Thesaurus'), and Notes. A large blue banner in the center-right area says 'Now 1 scheme'. At the bottom, there are buttons for 'ResView' and 'Code'.



The screenshot shows the VocBench application interface. The top navigation bar includes links for 'About VocBench', 'Current project: PLDN_Thesaurus' (with a refresh icon), and 'Global Data Management'. The main menu tabs are 'VocBench', 'Projects', 'Data' (selected), 'Metadata', 'SPARQL', and 'Tools'. On the left, a sidebar titled 'Concept' (highlighted with a red box) lists various categories and their sub-items, such as Applications, Assumptions, Network effect, Non-unique naming, Open World Assumption, Schema as you go, Best Practices, and Data. Under 'Best Practices', 'Data on The Web Best Practices (en)' is selected and highlighted with a blue background. The main workspace displays two projects: 'PLDN Thesaurus (en), PLDN Thesaur...' and 'Data on The Web Best Practices (en)'. The right-hand panel for the second project shows its properties: 'Types' (including rdf:type and skos:Concept), 'Top Concept of', 'Schemes' (skos:inScheme), 'Broaders' (skos:broader), and 'WDOO Recommendations for WDOO Annotations (en)'. A large blue banner in the center-right area of the workspace says 'Now 468 concepts and 22 top concepts'. The bottom of the screen features a search bar and navigation buttons for 'ResView', 'TermView', and 'Code'.



PLDN Thesaurus – Top concepts (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

S BegrippenXL: archixl: Alphabeti X + - ☰ ×

https://www.begrippenxl.nl/pldn/en/index/?clang=nl

... ☆ ☁ ⚡ ⌂

BegrippenXL

Home About Feedback Help | in Nederlands auf Deutsch

PLDN thesaurus

Content language Dutch ▾

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

Fully bilingual

09:04
30-9-2020



PLDN Thesaurus – Current situation (30/9)

***We have all the puzzle pieces now, but we
“only” have to put the puzzle together***





PLDN Thesaurus – Next steps?

PLDN Thesaurus Working Group

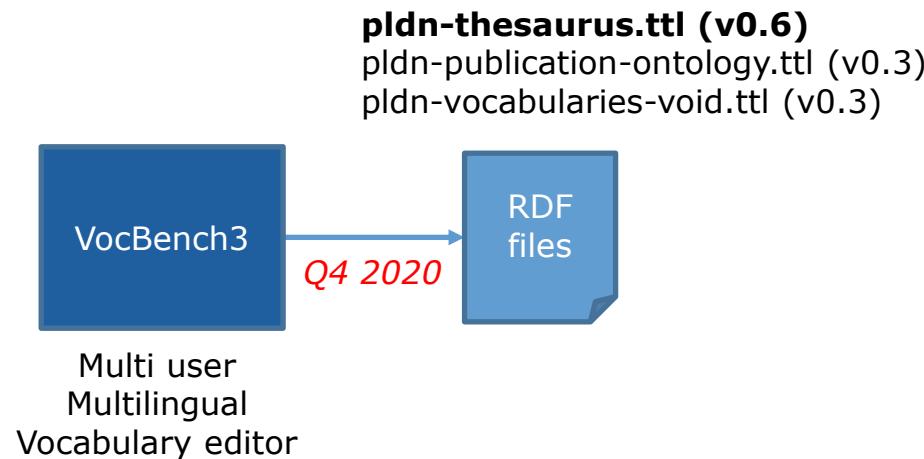
- Pieter van Everdingen
- Hans de Rond
- Pano Maria
- Jan Willem van Veen

New participants can apply via :

- info@pldn.nl



PLDN Thesaurus – Next steps?



PLDN Thesaurus Working Group

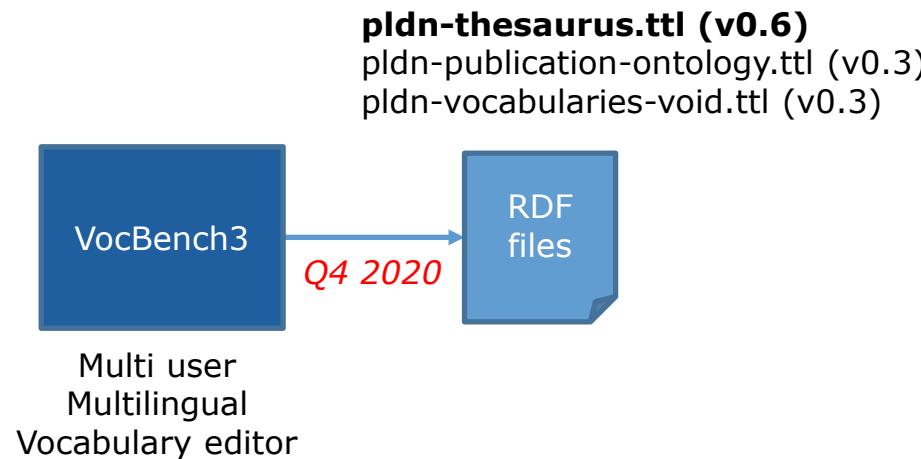
- Add editors
- Add reviewers

First activities:

- Agree upon work process
- Agree upon documentation process
- Agree upon naming convention
- Agree upon URI strategy
- Review first thesaurus version
- Process feedback in the next version of the thesaurus



PLDN Thesaurus – Next steps?



PLDN Thesaurus Working Group

- Add editors
- **Add reviewers**

First activities:

- **Agree upon work process**
- Agree upon documentation process
- Agree upon naming convention
- Agree upon URI strategy
- **Review first thesaurus version**
- Process feedback in the next version of the thesaurus



Platform Linked
Data Nederland

Work process – Reqs & issues on Github?

Screenshot of a GitHub organization page for "Platform Linked Data Nederland".

The top navigation bar includes a search bar, a repository icon, and links for Pull requests, Issues, Marketplace, and Explore. On the far right are icons for notifications, a plus sign, and user profile.

The organization's logo, a blue stylized snowflake or network icon, is displayed next to the name "Platform Linked Data Nederland".

The main navigation bar below shows tabs for Repositories (selected), Packages, People, and Projects.

The "Repositories" section displays a large icon of a smartphone and the text "This organization has no public repositories.".

The "People" section contains a message: "This organization has no public members. You must be a member to see who's a part of this organization." with a "People" link and a right-pointing arrow.

A "Report abuse" link is located at the bottom of the page.



Platform Linked
Data Nederland

Work process – Capture design decisions

The screenshot shows the homepage of the Platform Linked Data Nederland (PLDN) website. The header features the PLDN logo, the tagline "Innovatie door betekenisvol verbinden", a search bar with placeholder "Wat zoekt u?", and a navigation menu with links to "platform", "linked data", "experts", and "doe mee". The main content area displays the title "PLDN Thesaurus Best Practices" and the subtitle "PLDN Werkdocument 27 september 2020". Below this, there are sections for "Deze versie:", "Redacteur:", and "Reviewers:", each with a corresponding URL or name.

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

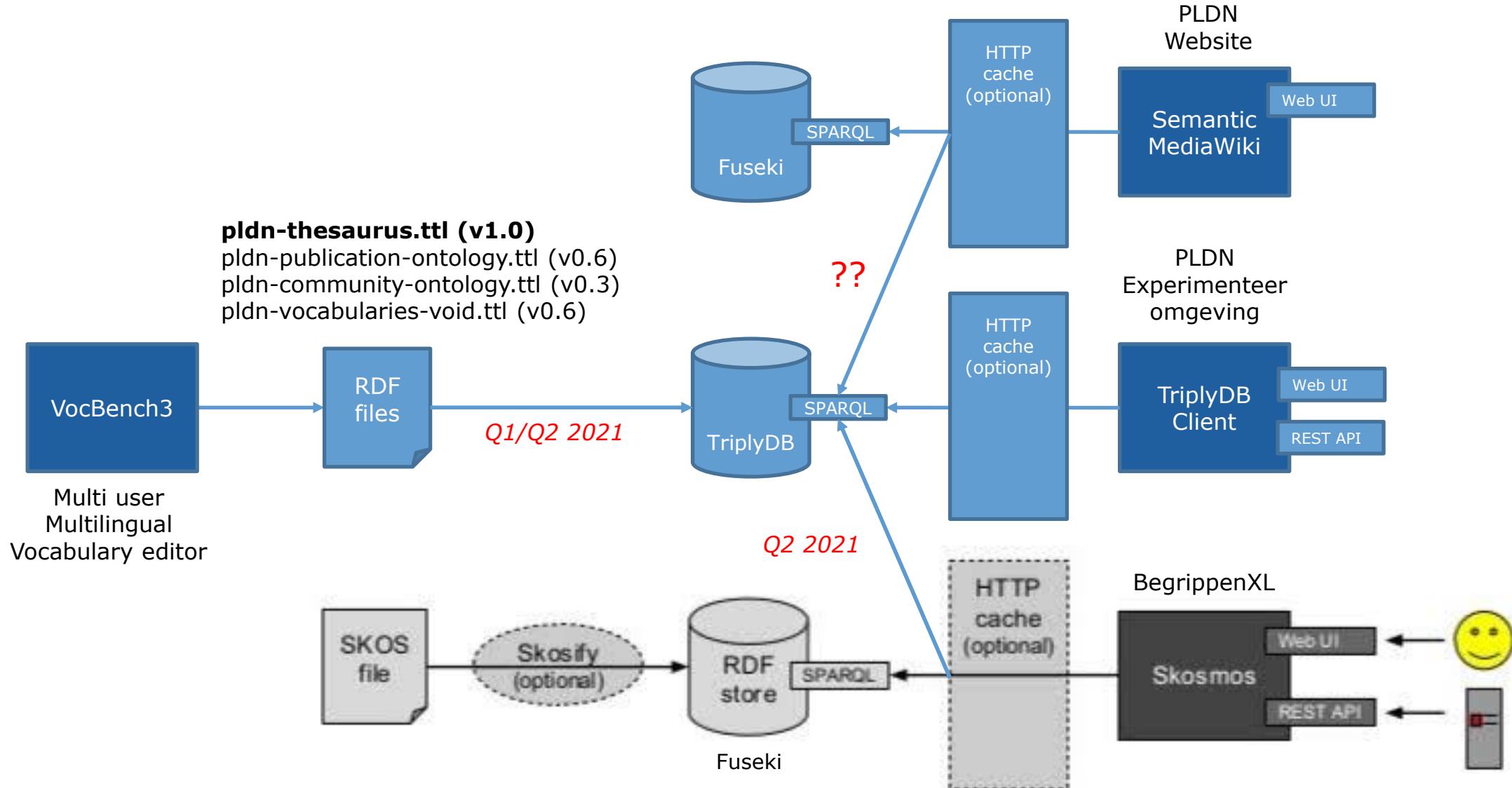
Reviewers:

Hans de Rond, Bureau De Rond
Pano Maria, Skemu



Platform Linked
Data Nederland

PLDN Thesaurus – Next steps?





PLDN Thesaurus – Take aways

- Domain knowledge, domain knowledge, domain knowledge...
- Collect as many concepts as possible (without boundary)
(this can help to see patterns more easily)
- Select a minimal set to get to the core concepts
- If a structure doesn't work, start all over again
(don't continue endlessly, if you're not satisfied)
- Only start using a tool, once you are satisfied
- Listen to the seasoned thesaurus pros in our field
- It is mostly a lot of manual work ... (human choices)





Platform Linked
Data Nederland

PLDN communication channels

BIG DATA
VALUE CENTER
OPEN INNOVATION CENTER

ECP

Platform voor de
InformatieSamenleving

Contact

Pieter van Everdingen

info@pldn.nl



Website

www.pldn.nl

Lab environment

www.data.pldn.nl

LinkedIn-group

www.linkedin.com/groups/466278

Twitter

[@linkeddata_nl](https://twitter.com/linkeddata_nl)
hashtag #LinkedDataNL

Newsletter

www.pldn.nl/wiki/Nieuwsbrieven



Platform Linked
Data Nederland

BACKUP SLIDES



PLDN Vocabularies – URI-strategy

- <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>



Each concept group has at least a few dozen and up to a few hundred terms (current estimate)



Innovatie door betekenisvol verbinden

Wat zoekt u?

platform linked data experts doe mee

experts expertise organisaties blogs

Expertises **Old version (now)**

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	D meer	L meer
• 303 redirect	• Data replicatie	• LIMES
• 3D geo-informatie	• Data schema's	• LOD Cloud
4	• Data science	• LOD Cloud Knowledge Graph vs COVID-19
• 4 Linked Data principes	• Data scraping	• LOD Laundromat
5	• Data services	• LOD2 Stack
	• Data sets	• LOTUS
	• Data shapes	• 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

Vocabulary information

PREFERRED TERM [pldn thesaurus](#)

Old version (now)

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