



Network of Terms

Bringing links to your data!

Enno Meijers - CTO Dutch Digital Heritage Network (NDE)

enno.meijers@kb.nl | <https://mastodon.social/@ennomeijers>

SWIB - 12 September 2023



dutch digital
heritage
network



Introduction to the NDE programme

The Dutch Digital Heritage Network (NDE)

The Dutch Digital Heritage Network (NDE) aims **at increasing the social value of the cultural heritage information** maintained by libraries, archives, museums and other cultural institutions.

The NDE strategy starts from the **end user perspective** and encourages institutions to provide digital heritage information that is more **visible, usable and sustainable**.

The NDE program is about building strong **cross sector networks** on the level of **expertise and information**. **Linked Data** is regarded as one of the enabling technologies.



Networks of cultural institutions



Archieven



Bibliotheken



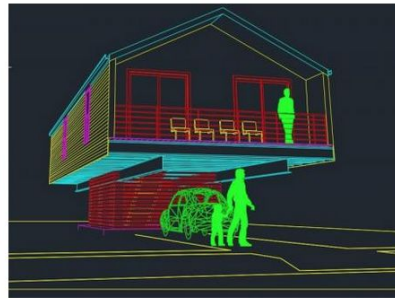
Musea



Media en AV



Digital humanities



Design en digitale cultuur



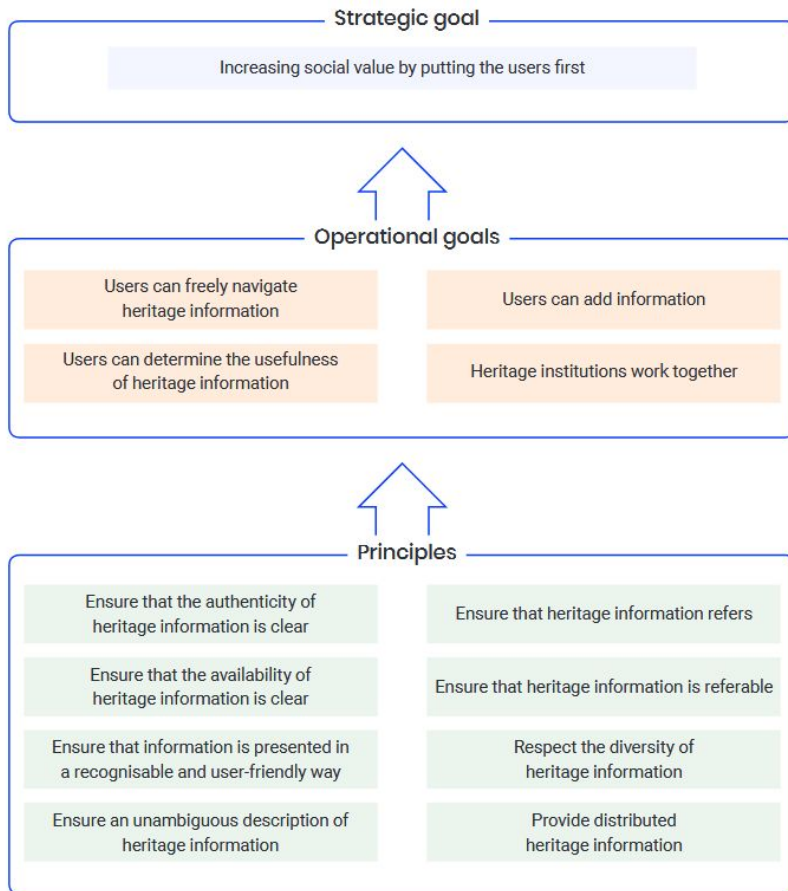
Podiumkunst.net



Netwerk Archieven Design en Digitale Cultuur



Cross domain reference architecture (DERA)



Overview of the NDE program

Visible

Service providers



Usage profiles



Campaign & channel



Rights & Usage

Usable

Infrastructure providers



Data & Terminologie Sources (LÖD)



Registries



Aggregators



Knowledge Graph

Sustainable

Source providers



PID



Preservation Policy & Certification



Cost model Preservation



Index Preservation Services

OVERALL



Supporting Network



Training & Education



Body of Knowledge



Services Toolbox



Service implementation & management



The background is a solid blue color. It features several white geometric patterns consisting of radiating lines. In the top-left corner, there is a circular pattern with a horizontal line through its center. In the top-right corner, there is a large, dense fan-shaped pattern. In the middle-left, there is a smaller fan-shaped pattern. In the bottom-left, there is another fan-shaped pattern. At the bottom center, there is a small fan-shaped pattern. In the bottom-right, there is a small circular pattern with a horizontal line through its center. All these patterns are connected to the edges of the frame by thin white lines.

Infrastructure for digital heritage

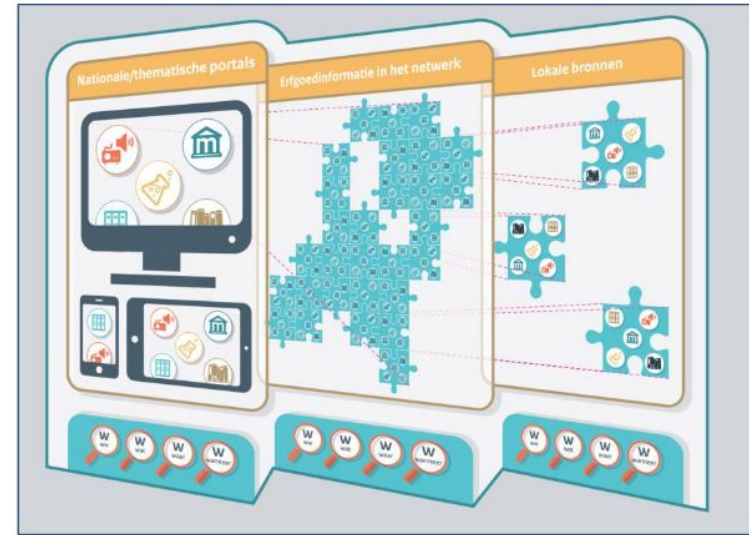
Infrastructure for digital heritage information

Rethink the architecture:

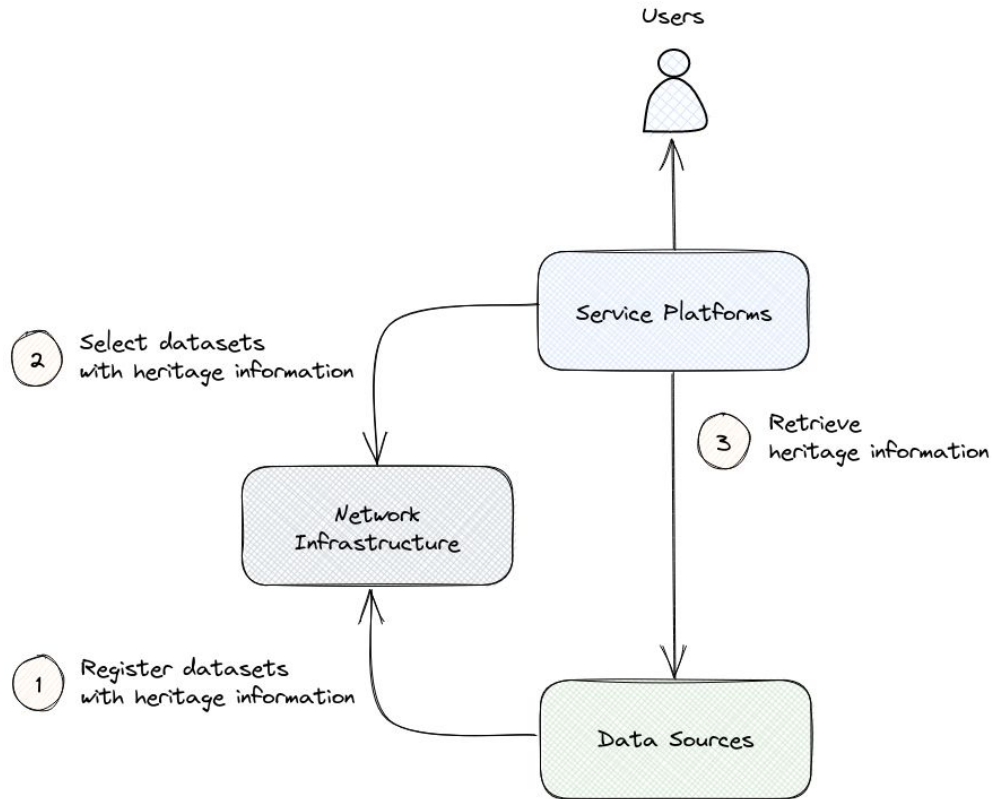
- maximize the usability of data at the source
- refer to data instead of copying
- build portals as dynamic views based on a common, interconnected data layer
- minimize the intermediate layers
- improve visibility on the web in general

Apply:

- Linked Data / FAIR principles
- 'web-centric' technologies (HTTP, RDF, Web APIs)
- decentralized technologies where possible



Service platforms – generic and dynamic services



BAROK 1600



Mars en Venus betragt door Vulcanus
Wtewael, Joachim, 1601



De val van Phaëthon
Rottenhammer, Hans (I), 1604



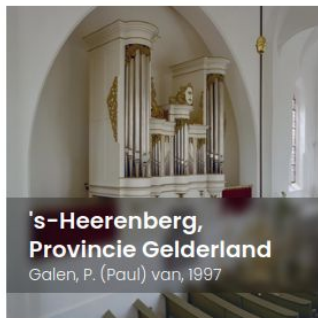
Boslandschap met jagers en waarzegster
Govaerts, Abraham, 1612



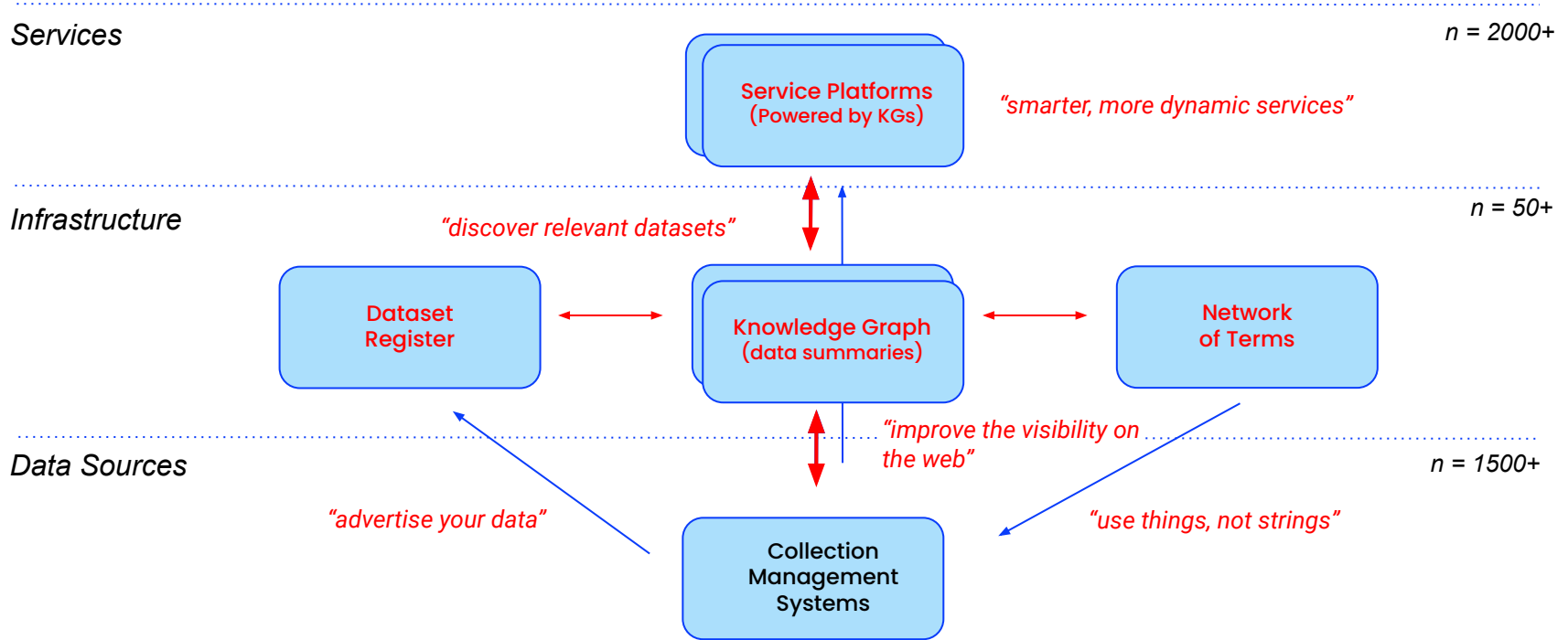
Portret van een officier
Ravesteyn, Jan van, 1612



PROVINCIE GELDERLAND



Roadmap National Infrastructure for Digital Heritage





Network of Terms

Using Linked Data in a network...



[A Pair of Leather Clogs](#)
(Van Gogh Museum)



[Garden at Arles](#)
(Kunstmuseum)



[National Library](#)

Dataset Registry

schema:creator

rico:isCreatorOf

schema:creator

rico:isCreatorOf

schema:about

schema:subjectOf

Knowledge Graph



Vincent van Gogh
<https://data.rkd.nl/artists/32439>
([RKD Artists](#))

Network of Terms



Available sources for linking through the Network of Terms

National Institutes

- Cultural Heritage Agency (thesauri)
- National Library (thesauri, persons/org.)
- Sound & Vision Institute (thesauri, persons/org.)
- Music library of The Netherlands (thesauri, persons/org.)
- Performing Arts Network (thesauri, persons/org.)
- Netherlands Institute for Art History (persons)
- Nationaal Museum of World Cultures (thesauri)
- Indonesia Remembrance Centre (thesauri)
- LGBTI Heritage Organisation (thesauri)
- Second World War Documentation (thesauri, persons)

In numbers:

- Organisations: 18
- Individual terminology sources: 25
- Search queries: 40

International Institutes

- Getty Research Institute (AAT)
- Geonames (NL, BE, DE)
- Wikidata (entities, persons, places, streets)
- European Union (EuroVoc)
- Henri van de Waal Foundation (Iconclass)

Regional Institutes

- Erfgoed Brabant (buildings)
- Gouda Time Machine (streets)
- Cultural Institutions Amsterdam (streets)

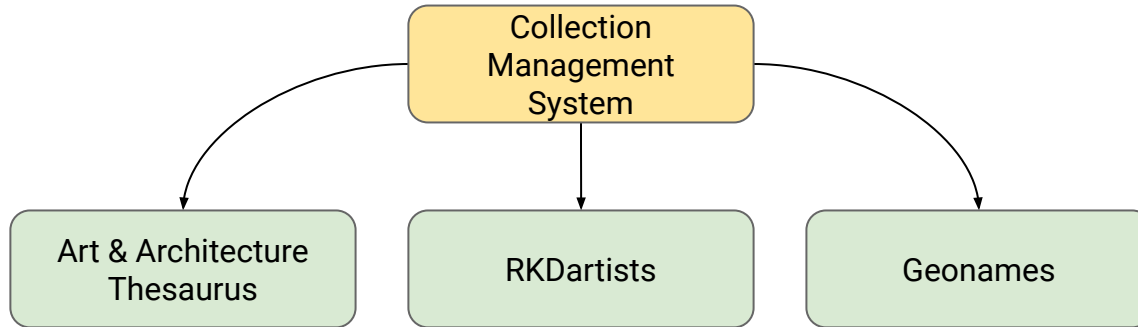
Default direct approach

Use Case:

*Find relevant Term URI(s) for concepts, persons, places,... when describing resources in **my own** system.*

Challenges:

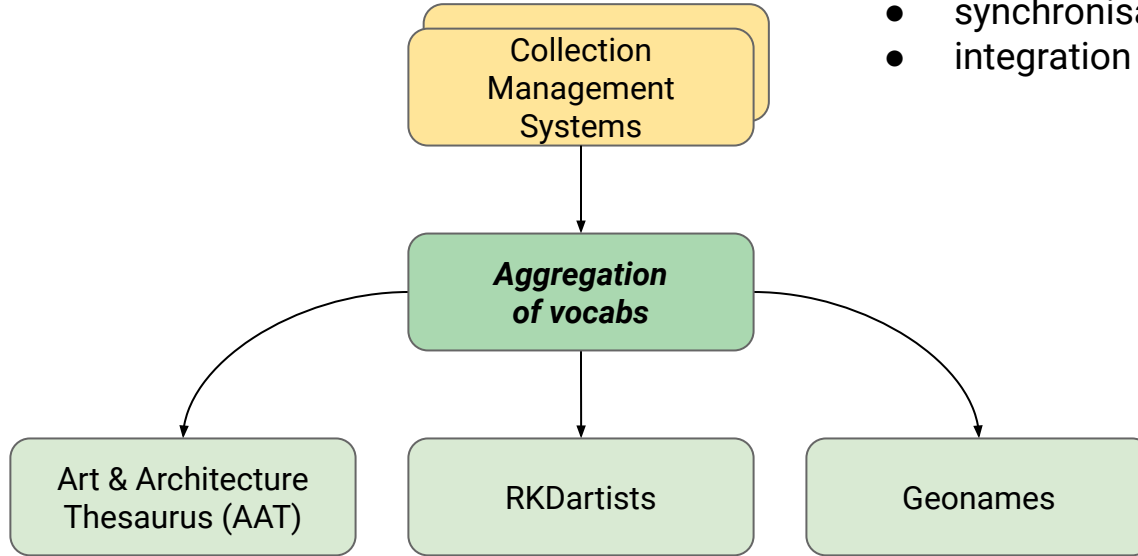
- harvest sources or query live apis?
- deal with multiple protocols
- deal with multiple data models
- complexity and maintenance issues



'From string to thing': "Van Gogh" → <https://data.rkd.nl/artists/32439>



Default platform approach



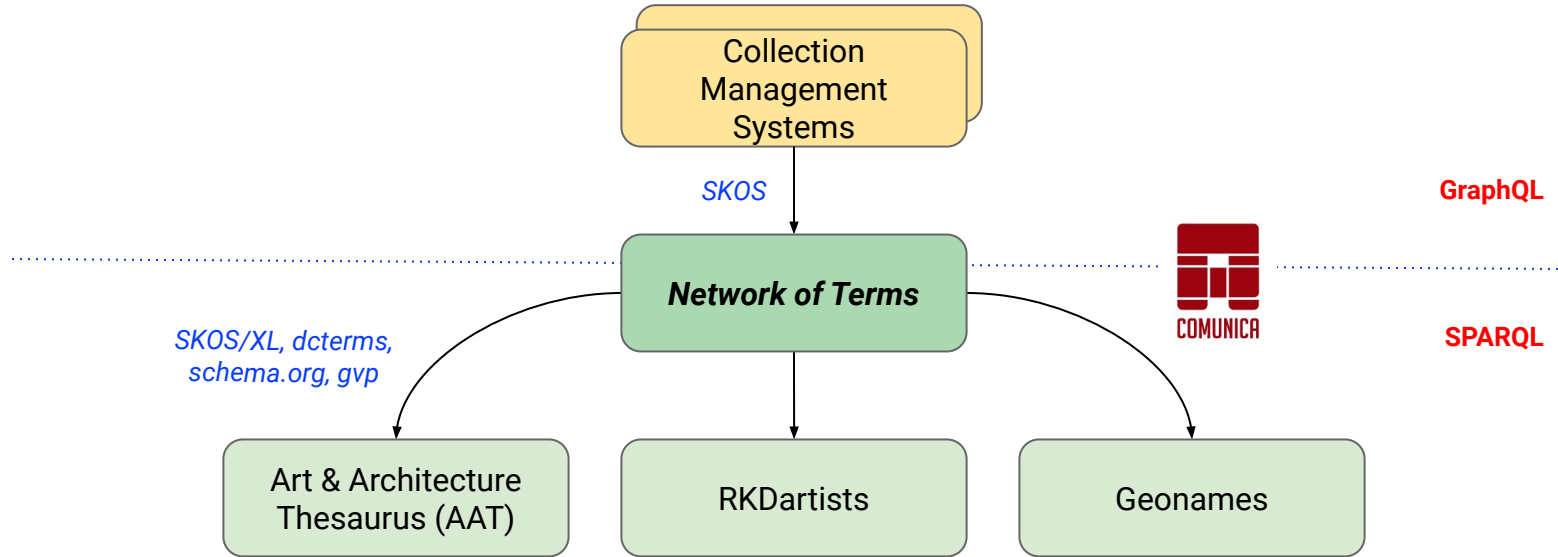
Some challenges:

- deal with multiple data models
- synchronisation with source data
- integration with other systems?



Network of Terms approach

Harmonisation of data models and protocols



Searching shared thesauri, classification systems, and reference lists

for collection administrators | [for administrators of terminology resources](#) | for heritage software developers

Search terms **Look up term**

Search words

Terminology sources

Live demo...

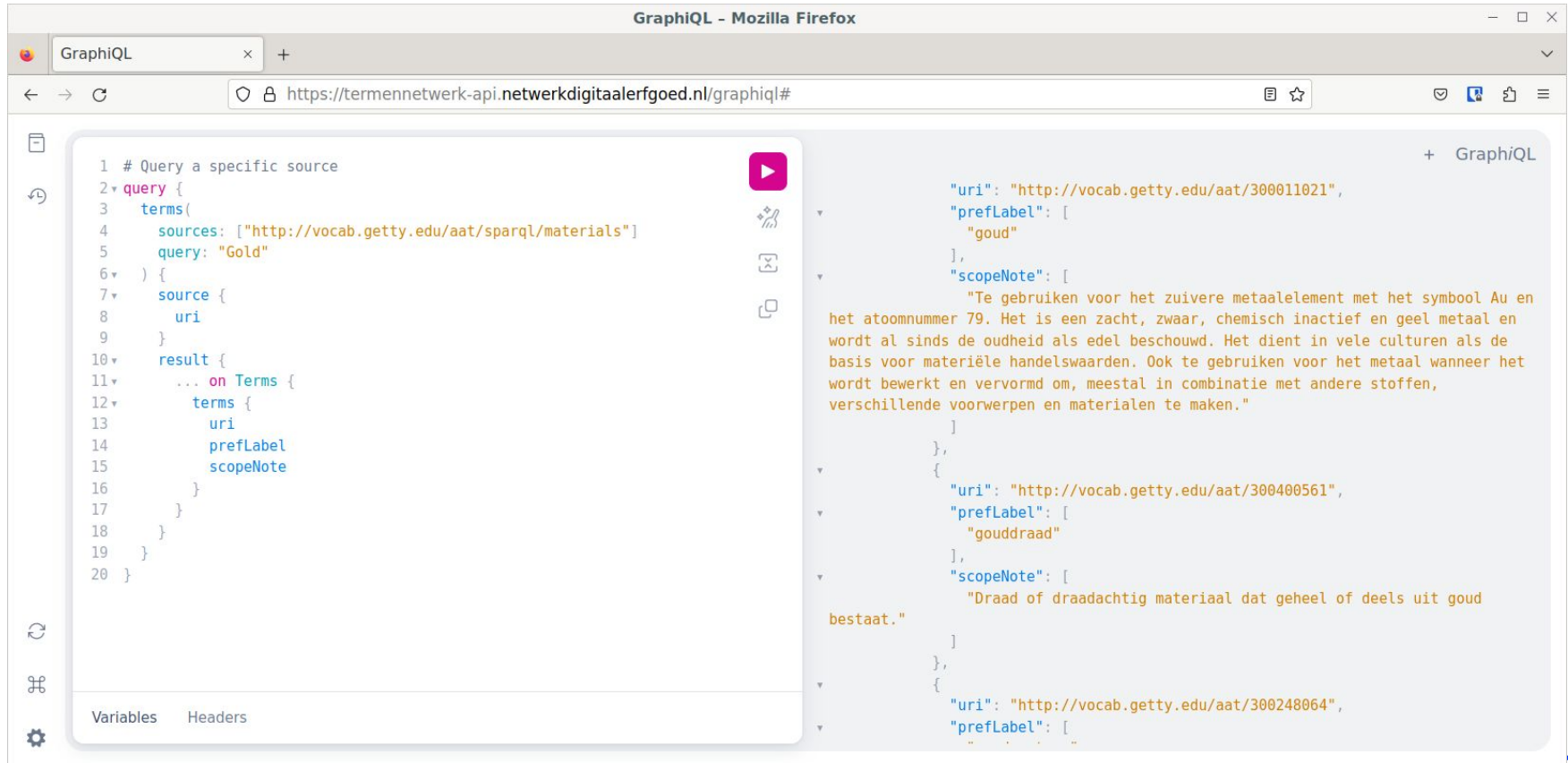
Art & Architecture Thesaurus - materialen (AAT - materialen)
... materialen in architectuur-, kunst- en cultuurhistorische collecties
Getty Research Institute
43 terms found (in 1252 ms)

<metaal naar kwaliteit>

Broader term



But it is really about the API...



The screenshot shows a web browser window titled "GraphiQL - Mozilla Firefox" with the address bar displaying "https://termennetwerk-api.netwerkdigitaalergoed.nl/graphiql#". The main content area is a GraphQL IDE with a query editor on the left and a response viewer on the right.

```
1 # Query a specific source
2 query {
3   terms(
4     sources: ["http://vocab.getty.edu/aat/sparql/materials"]
5     query: "Gold"
6   ) {
7     source {
8       uri
9     }
10    result {
11      ... on Terms {
12        terms {
13          uri
14          prefLabel
15          scopeNote
16        }
17      }
18    }
19  }
20 }
```

The response viewer shows a JSON array of objects, each representing a term. The first object is for "goud" (gold) and the second is for "gouddraad" (gold wire). The response is partially collapsed in the image.

```
{
  "uri": "http://vocab.getty.edu/aat/300011021",
  "prefLabel": [
    "goud"
  ],
  "scopeNote": [
    "Te gebruiken voor het zuivere metaalelement met het symbool Au en het atoomnummer 79. Het is een zacht, zwaar, chemisch inactief en geel metaal en wordt al sinds de oudheid als edel beschouwd. Het dient in vele culturen als de basis voor materiële handelswaarden. Ook te gebruiken voor het metaal wanneer het wordt bewerkt en vervormd om, meestal in combinatie met andere stoffen, verschillende voorwerpen en materialen te maken."
  ]
},
{
  "uri": "http://vocab.getty.edu/aat/300400561",
  "prefLabel": [
    "gouddraad"
  ],
  "scopeNote": [
    "Draad of draadachtig materiaal dat geheel of deels uit goud bestaat."
  ]
},
{
  "uri": "http://vocab.getty.edu/aat/300248064",
  "prefLabel": [
    ..
  ]
}
```



Example of a SPARQL CONSTRUCT query used for transformation to SKOS

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
```

```
PREFIX void: <http://rdfs.org/ns/void#>
```

```
CONSTRUCT {
```

```
  ?uri a skos:Concept ;  
  skos:prefLabel ?rdfs_label ;  
  skos:altLabel ?schema_name ;  
  skos:altLabel ?schema_alternateName ;  
  skos:scopeNote ?schema_description .
```



Map all types to skos:Concept for easy access

```
}
```

```
WHERE {
```

```
  ?uri foaf:isPrimaryTopicOf/void:inDataset <http://data.bibliotheken.nl/id/dataset/stcn> ;  
  schema:additionalType <http://www.productontology.org/id/Printer_%28publishing%29> ; # Select printers.  
  rdfs:label ?rdfs_label .
```

```
  ?uri ?predicate ?label .
```

```
VALUES ?predicate { rdfs:label schema:name schema:alternateName }
```

```
?label <bif:contains> ?virtuosoQuery .
```



some "magic" for full text search depending on SPARQL-endpoint type

```
OPTIONAL { ?uri schema:name ?schema_name }
```

```
OPTIONAL { ?uri schema:alternateName ?schema_alternateName }
```

```
OPTIONAL { ?uri schema:description ?schema_description }
```

```
}
```

```
LIMIT 1000
```

mapping

selection



Reconciliation Service API



TABLE OF CONTENTS

Abstract

Status of This Document

1. Introduction

- 1.1 Data Matching on the Web
- 1.2 History of the Reconciliation API
- 1.3 External Resources
- 1.4 Versions
 - 1.4.1 0.1
 - 1.4.2 0.2 (This Version)
- 1.5 Conformance

2. Core Concepts

- 2.1 Entities
- 2.2 Types
- 2.3 Properties
- 2.4 Property Values
- 2.5 Identifier and Schema Spaces

3. Service Definition

- 3.1 Service Manifest
- 3.2 HTTP(S) Access

Reconciliation Service API v0.2

A protocol for data matching on the Web

Final Community Group Report 10 April 2023

This version:

<https://www.w3.org/community/reports/reconciliation/CG-FINAL-specs-0.2-20230410/>

Latest published version:

<https://www.w3.org/community/reports/reconciliation/CG-FINAL-specs-0.2-20230410/>

Latest editor's draft:

<https://reconciliation-api.github.io/specs/draft/>

Editors:

[Antonin Delpuch](#)  (University of Oxford)
[Adrian Pohl](#)  (Hochschulbibliothekszentrum NRW)
[Fabian Steeg](#)  (Hochschulbibliothekszentrum NRW)
[Thad Guidry Sr.](#) 
[Osma Suominen](#)  (National Library of Finland)

Feedback:

[GitHub reconciliation-api/specs](#) (pull requests, new issue, open issues)
public-reconciliation@w3.org with subject line [specs-0.2] ~ message topic ~ (archives)

Copyright © 2023 the Contributors to the Reconciliation Service API v0.2 Specification, published by the [Entity Reconciliation Community Group](#) under the [W3C Community Final Specification Agreement \(FSA\)](#). A human-readable [summary](#) is available.

Abstract

This document describes the reconciliation service API, a protocol edited by the [W3C Entity Reconciliation Community Group](#). It is intended as a comprehensive and definitive specification of



Add links to existing data using OpenRefine:

10000 rows

Show as: rows records Show: 5 10 25 50 rows

All	s	name	birthDateStr	deathDateStr
☆		301. http://data.bibliotheken.nl/id/thes/p421470488 M. Keur <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Keur, M. (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Keur, M. v.d. (85) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item	1997	
☆		302. http://data.bibliotheken.nl/id/thes/p421670983 Dieuwer van Greevenbroek <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Greevenbroek, Dieuwer van (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item	1997	
☆		303. http://data.bibliotheken.nl/id/thes/p421673664 Julian de Bondt <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Bondt, Julian de (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		
☆		304. http://data.bibliotheken.nl/id/thes/p427868505 Anne Stijnen <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Stijnen, Anne (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		
☆		305. http://data.bibliotheken.nl/id/thes/p430969694 Marit Helwig <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Helwig, Marit (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		
☆		306. http://data.bibliotheken.nl/id/thes/p431161011 Lale Gül <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Gül, Lale (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		
☆		307. http://data.bibliotheken.nl/id/thes/p417553005 Elizabeth Visser <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Visser, Elizabeth (1908-1987) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Visser, Elizabeth (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Visser, Karin de (1975-) (89) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Sjamsoedin-Visser, Elizabeth Jacoba Margaretha (1942-) (76) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		
☆		308. http://data.bibliotheken.nl/id/thes/p419034587 Haryanti Frateur <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Frateur, Haryanti (1997-) (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Create new item		

Match this Cell Match All Identical Cells

Cancel

Sjamsoedin-Visser, Elizabeth Jacoba Margaretha (1942-)

Alternatieve labels
Sjamsoedin-Visser, Liesbeth • Elizabeth Jacoba Margaretha Sjamsoedin-Visser • Visser, L.

Every Network of Terms source gets a Reconciliation Service endpoint by default!



DIY – Network of Terms Tutorial

network-digitaal-erfgoed / network-of-terms-tutorial

Q Type 🔍 to search

<> Code Issues Pull requests Actions Projects Wiki Security Insights Settings

Configuring the catalog

Enno Meijers edited this page 3 weeks ago · 9 revisions

Edit New page

Introduction to the Catalog

The Catalog is the core component of the Network of Terms. Adding new terminology sources to the Network of Terms is only a matter of making changes to the Catalog configuration. **No changes of the code itself are necessary.** The Catalog only exists of files containing JSON-LD objects or SPARQL queries.

To make changes to the Catalog it is important to understand the structure and data model behind it.

The Catalog is part of the [network-of-terms repository](#) and is a separate package located in `packages/network-of-terms-catalog`. The Catalog has the following file structure:

```
* catalog/
* publishers.jsonld # definition of organizations that publish data sources
* datasets/
* <ds_name_n>.jsonld # dataset definition for each data source
* queries/
* search/
  * <qname_n>.rq # SPARQL queries for searching
* lookup/
  * <qname_n>.rq # SPARQL queries for lookups
```

Data Model

Pages 19

- Home
- Goal and purpose
- Current system integrations
- Main components
 - GraphQL API
 - Demonstrator
 - Reconciliation Service API
- Run your own Network of Terms instance
 - Running the GraphQL API
 - Running the Demonstrator
 - Running the Reconciliation Service API
- Catalog configuration
 - Making changes to the queries
 - Adding a new data source
 - Setup your own catalog

<https://github.com/network-digitaal-erfgoed/network-of-terms-tutorial/wiki>



So far we have found the following examples of that deploy full text search capabilities:

- **GraphDB - Lucene FTS plugin (deprecated)**

Used in Getty sparql endpoint serving AAT, GTN and others:

```
?uri luc:term ?query
```

`luc:term` : Brief, includes all terms (prefLabels and altLabels) and subject ID (default)

`luc:text` : Full, includes all terms, qualifiers, subject ID, and scope notes.

- **GraphDB - Lucene connector:**

If the Lucene connector is configured the available indexes can be listed in the following way:

```
PREFIX luc: <http://www.ontotext.com/connectors/lucene#>
SELECT ?cntUri ?cntStr {
  ?cntUri luc:listConnectors ?cntStr .
}
```



- **GraphDB - Elasticsearch connector:**

If the Elasticsearch connector is configured the available indexes indexes can be listed in the following way:

```
PREFIX elastic: <http://www.ontotext.com/connectors/elasticsearch#>
SELECT ?cntUri ?cntStr {
  ?cntUri elastic:listConnectors ?cntStr .
}
```



Searching by using one of the indexes is shown in this example from the [SemOpenAlex endpoint](#):

```
?search a elastic-index:semopenalex-authors ;
  elastic:query ?query ;
  elastic:entities ?uri .
?uri elastic:score ?score .
```



See for more info <https://graphdb.ontotext.com/documentation/10.0/graphdb-connectors.html>

- **Apache Jena Fuseki**

Example for using the Lucene implementation with a Fuseki endpoint:

```
(?uri ?score) text:query (<field1>...<fieldn> ?query 100)
```



Implementation

Integrations of Network of Terms API in other systems:

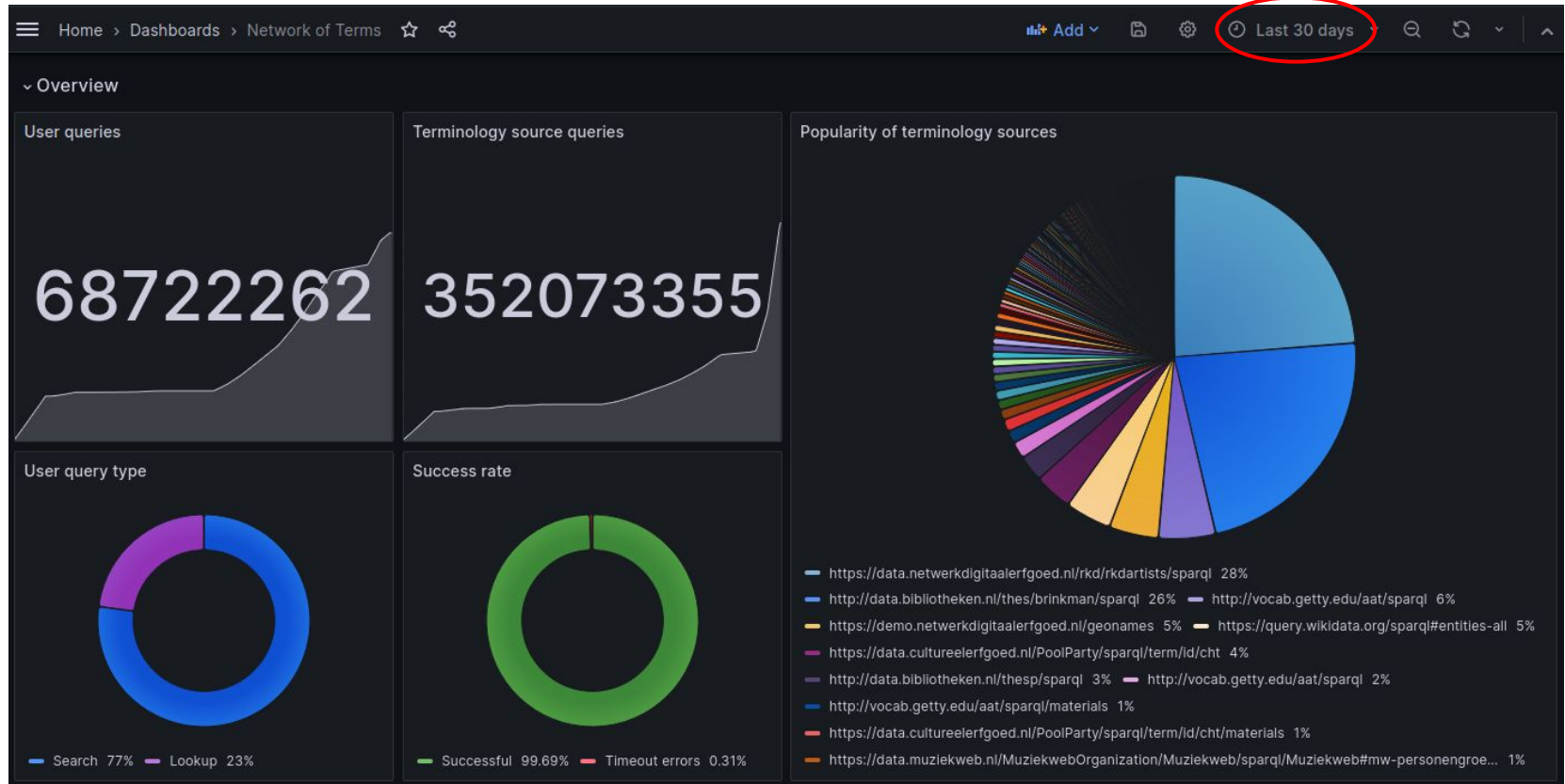
- Atlantis (DEVENTit)
- Memorix (Picturae)
- Axiell Collections (Axiell)
- Kleksi (Sofco)
- MuseumPlus (ZetCom - under discussion)
- Omeka-S (open source)
- Solid-CRS (open source - experimental)
- LDWizard - Heritage (open source - limited functionality)
- Mark Lindeman - VSCode Termennetwerk Extension (open source)

Source code (EURL-licensed) available at

<https://github.com/netwerk-digitaal-erfgoed/network-of-terms>



Usage over August '23



Takeaways:

- Federated querying of linked data works (for our use case)
 - Comunica.dev is a powerful framework for federated querying over Linked Data
 - Standardisation of full text search for SPARQL is needed
 - Network-of-Terms is a generic tool, it could fit your use case too
- => give it a try and share your ideas with us!





Thank you!

Enno.meijers@kb.nl | <https://mastodon.social/@ennomeijers>

Learn more at:

<https://github.com/netwerk-digitaal-erfgoed/network-of-terms-tutorial/wiki>

netwerkdigitaalerfgoed.nl



**dutch digital
heritage
network**