Managing changes in knowledge organization systems through linked data technologies: the case of UDC

Aida Slavic & Ana Vukadin
UDC Consortium, The Hague

Metadata Standards Matter
19 March 2025
National Library of Greece, Athens

Overview

01 **UDC in LD environment**
Supporting semantic integration

02 **Reality check**
Not all data can be linked

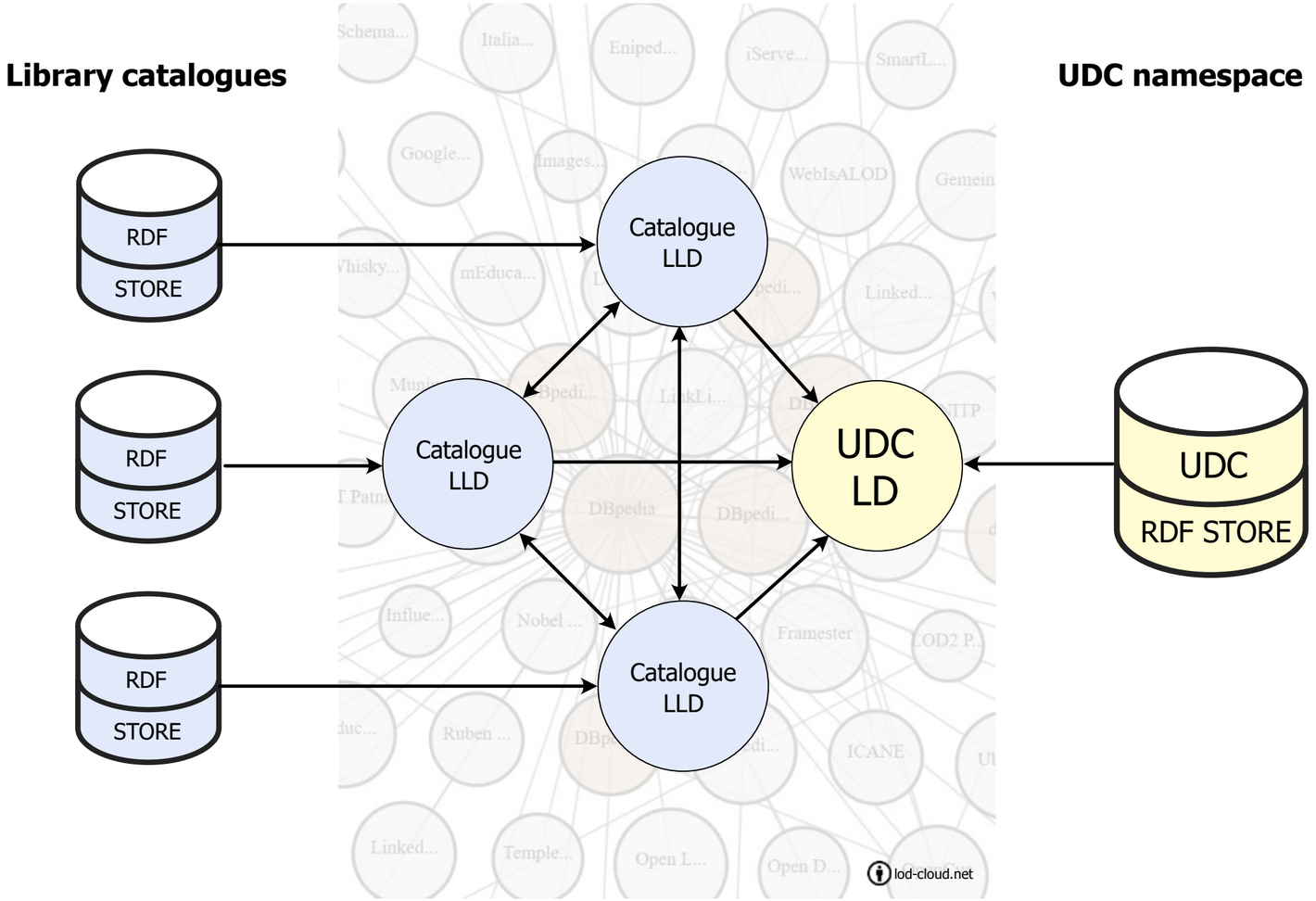
03 **UDC solution**
Look-up service

04 **Challenges**
Testing and implementation

05 **Conclusion**
Concluding remarks

UDC in LD environment

Supporting semantic integration



mapping to other KOS

multilingual access

contextualization

Reality check

Not all data can be linked

1

cancelled notations (concepts moved from one place to another)

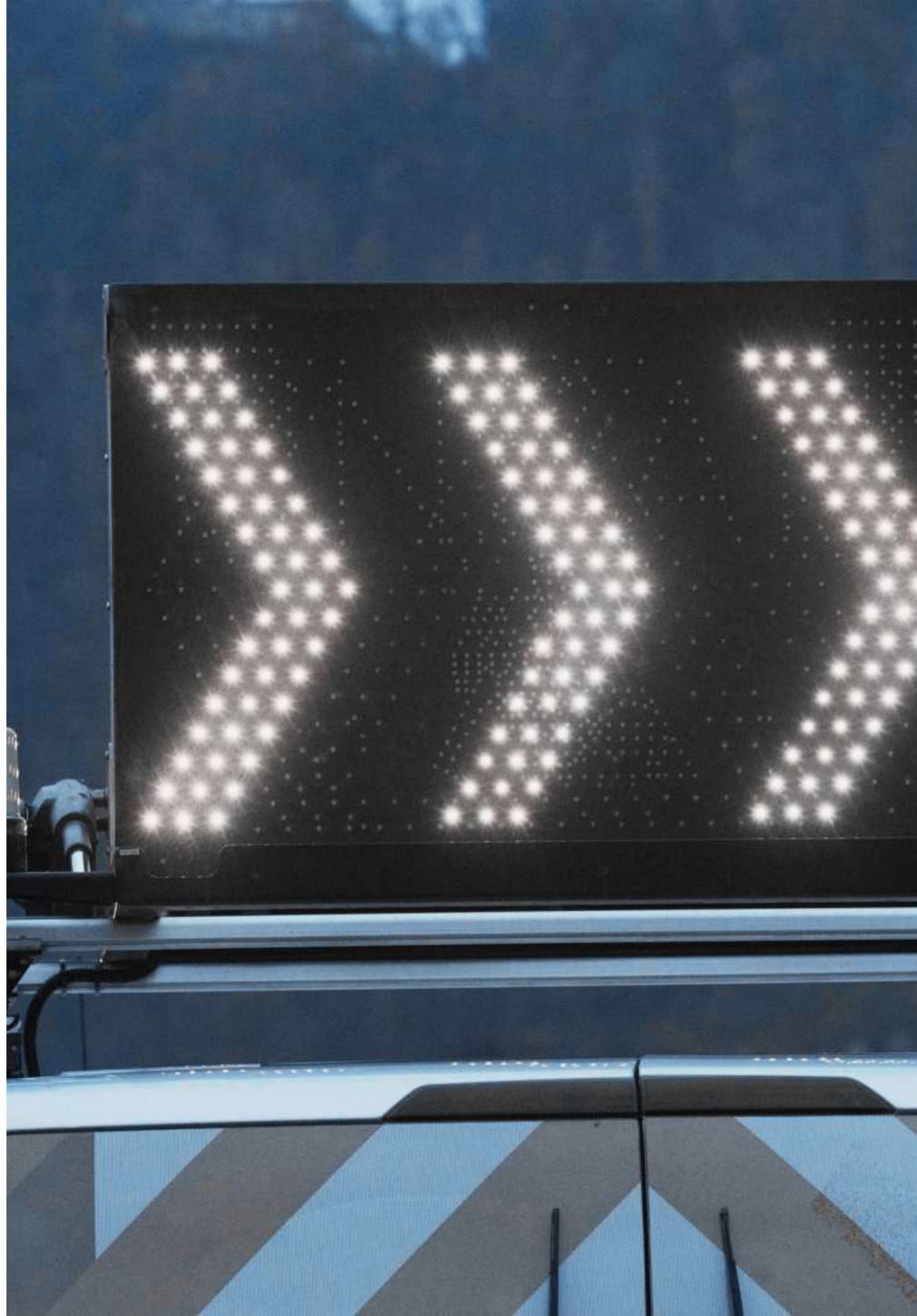
Reality check

Cancelled notations: since 1992

40% content changed

12.000 notations cancelled

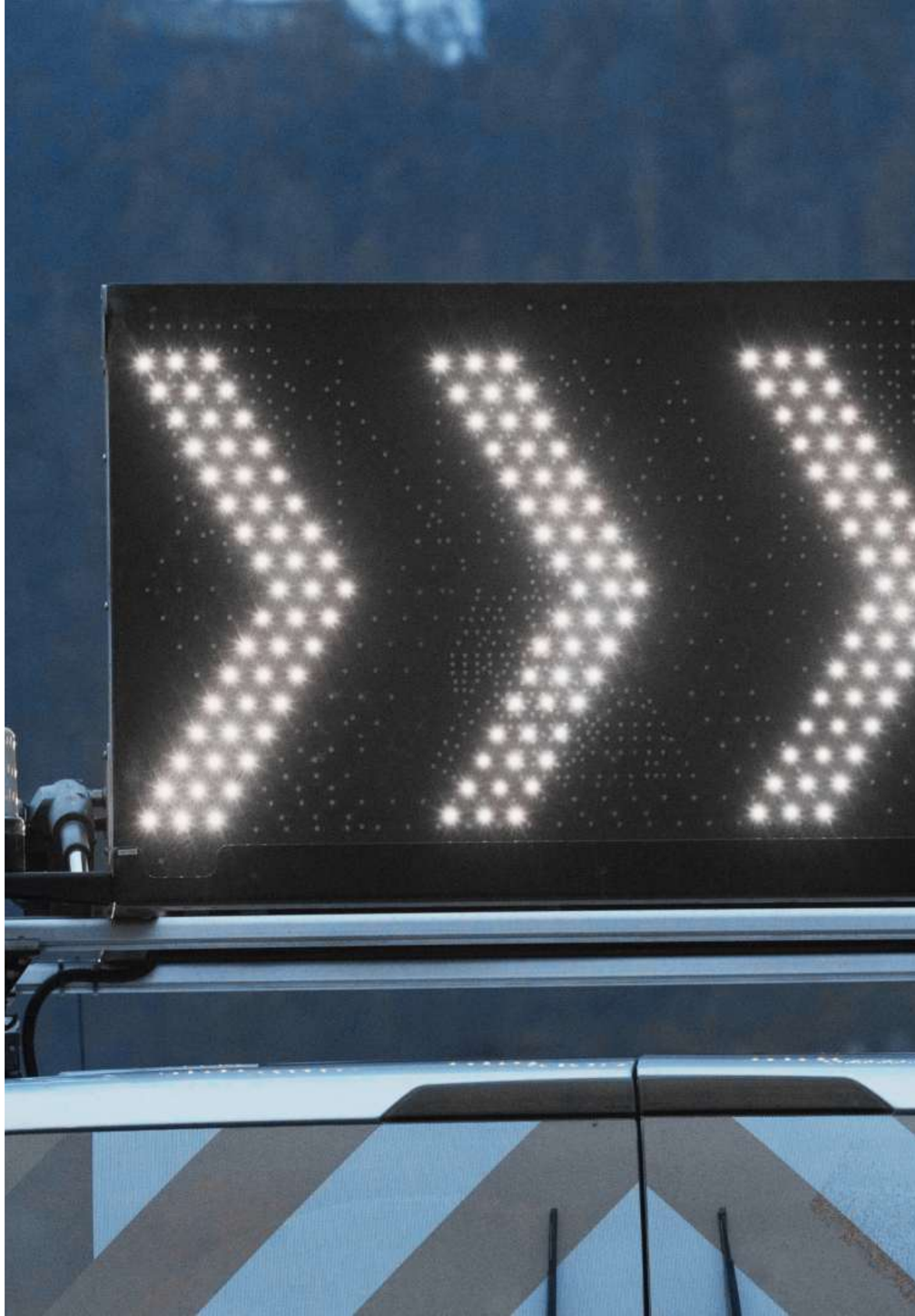
23.000 notations added



Reality check

Cancelled notations: example

968.0	**x [1994, E&C16] History of South Africa** → 94(680)



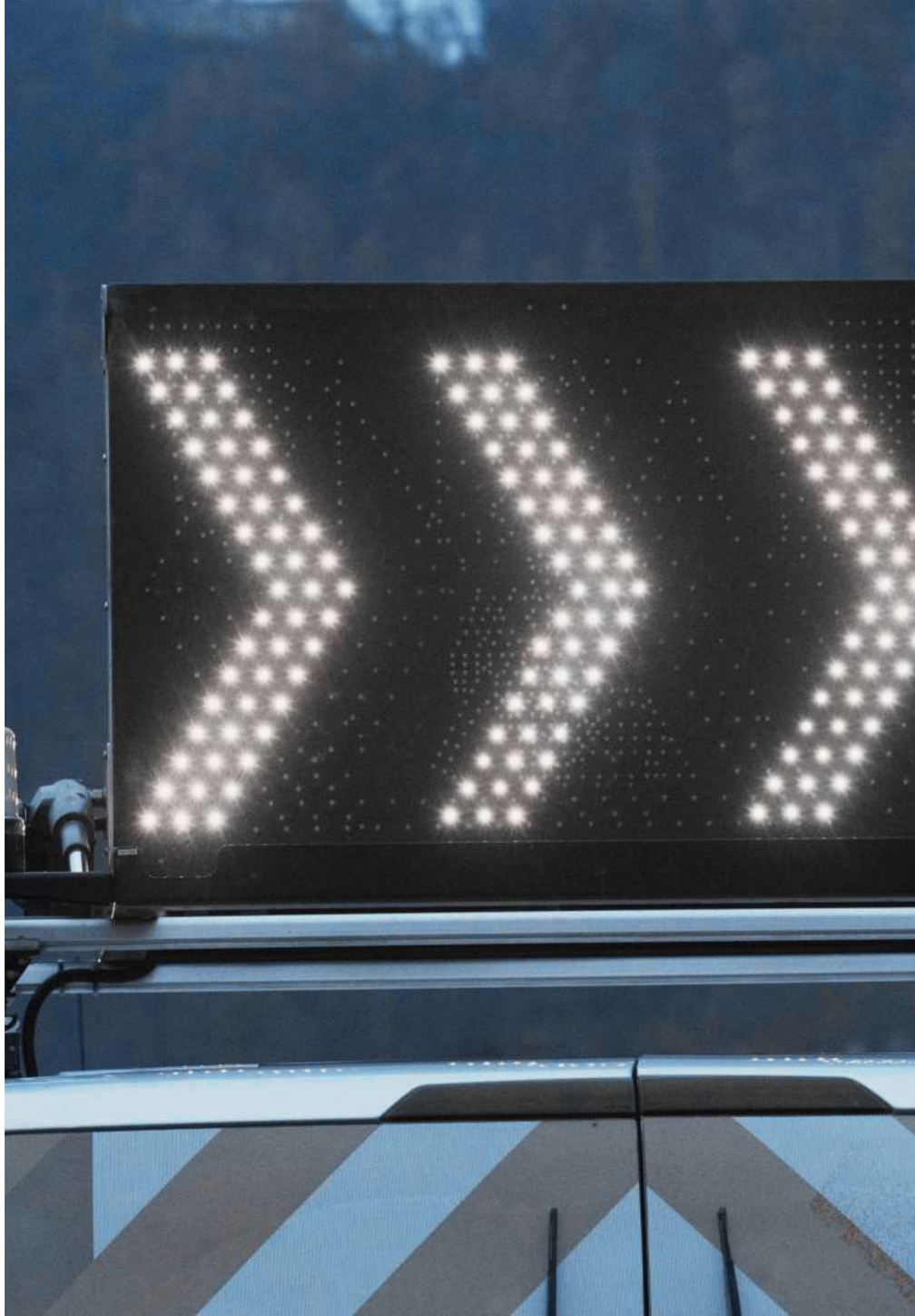
Reality check

Cancelled notations: example

HISTOIRE DE L'AFRIQUE DU SUD

AUTOR(ES):	Lugan, Bernard
PUBLICACAO:	Paris : Libr. Acad. Perrin, cop. 1990
DESCR. FISICA:	288, [8] p.
ISBN:	2-262-00847-7
ASSUNTOS:	África do Sul -- História
CDU:	968.0

Source: https://porbase.bnportugal.gov.pt

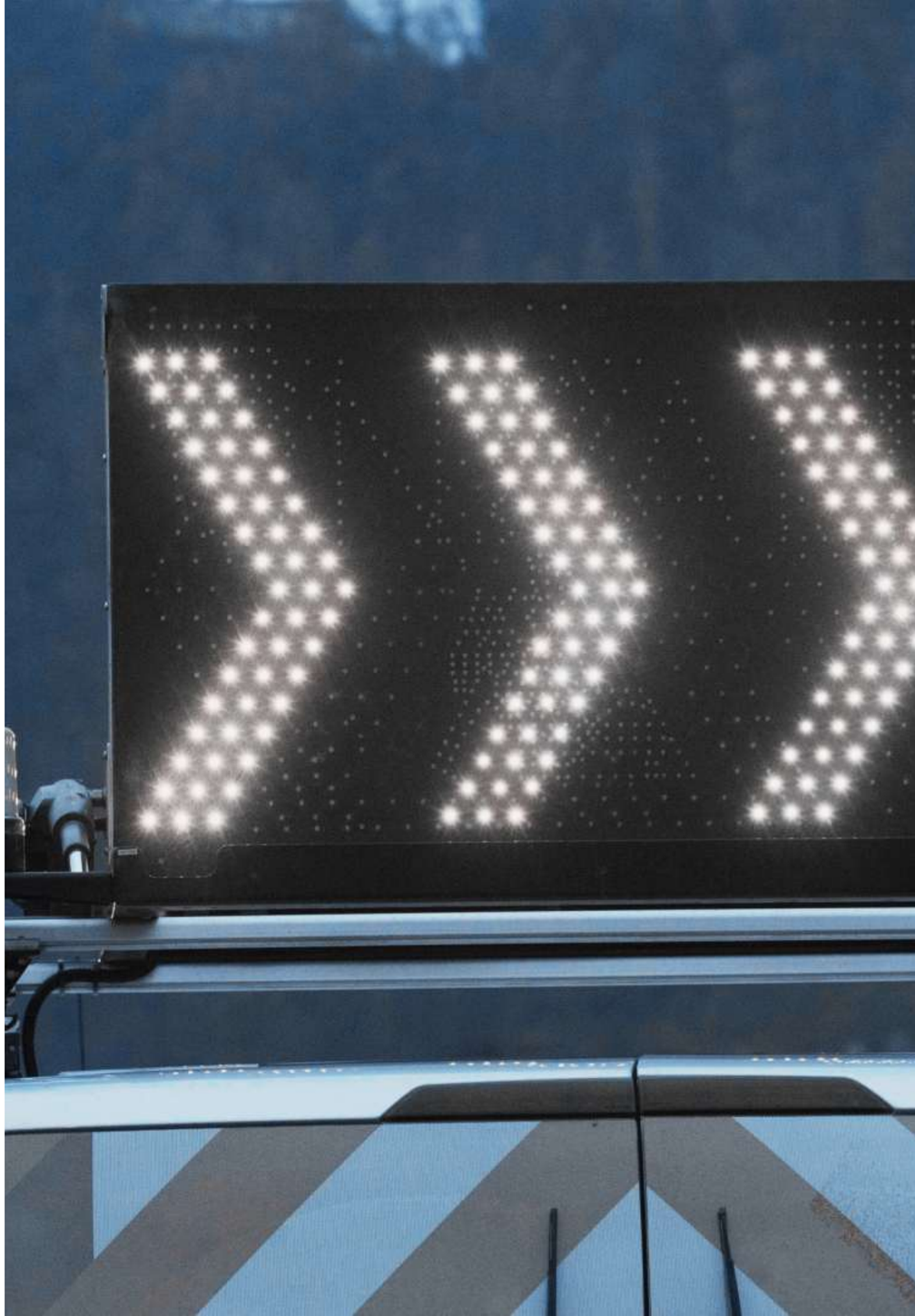


Reality check

Cancelled notations: example

```
<rdf:RDF>
  <rdf:Description>
    <dc:title>Histoire de l'Afrique du Sud</dc:title>
    <dc:creator>Lugan, Bernard</dc:creator>
    <dc:subject xsi:type="dcterms:UDC">968.0, por</dc:subject>
    <dc:description>Monografia</dc:description>
    <dc:publisher>Libr. Acad. Perrin</dc:publisher>
    <dc:publisher>Paris</dc:publisher>
    <dc:date>cop. 1990</dc:date>
    <dc:type>material textual, impresso</dc:type>
    <dc:format>288, [8] p.</dc:format>
    <dc:language xsi:type="dcterms:ISO 639-2">fre</dc:language>
  </rdf:Description>
</rdf:RDF>
```

Source: https://opendata.bnportugal.gov.pt/



Reality check

Not all data can be linked

1

cancelled notations (concepts moved from one place to another)

2

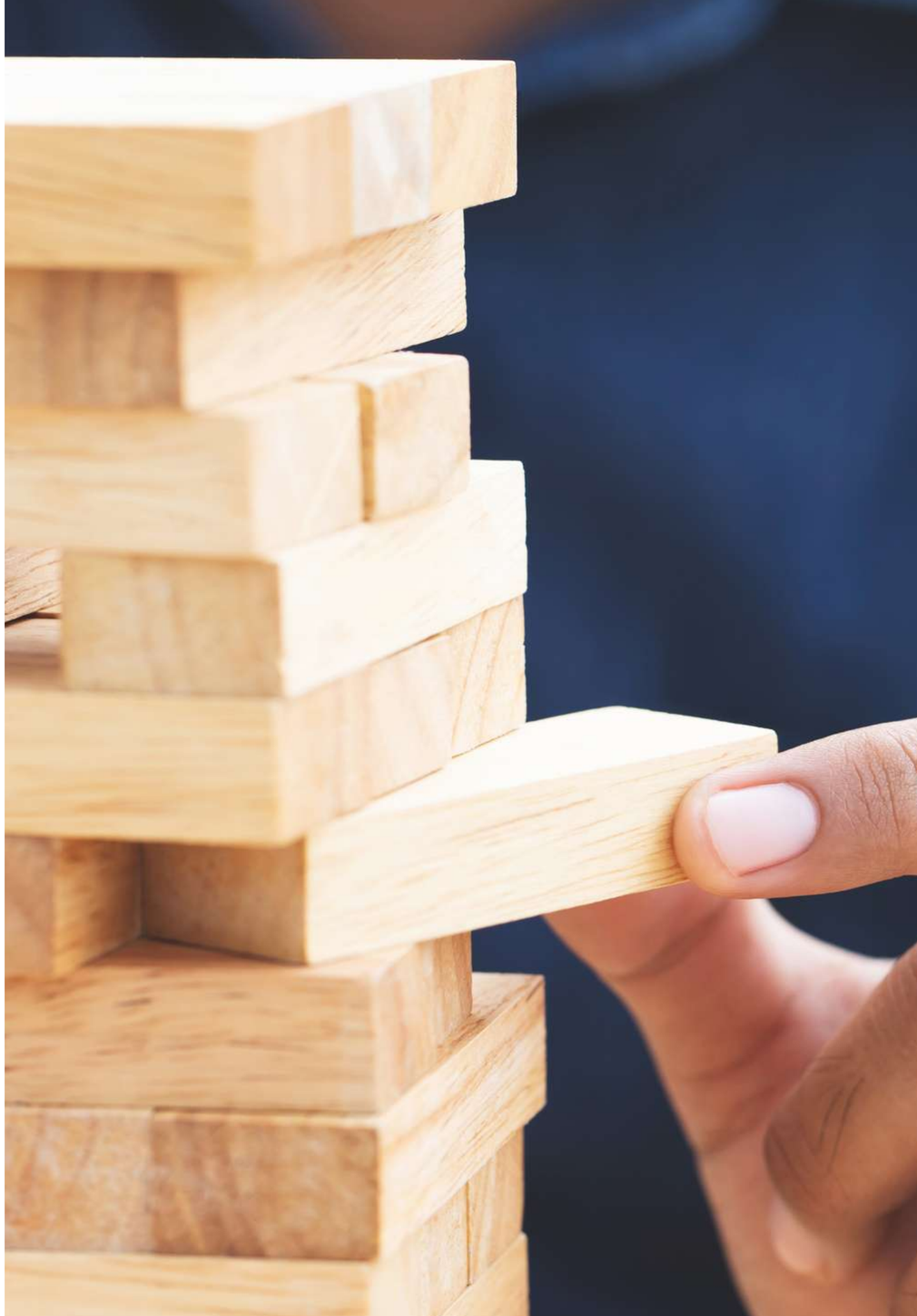
complex syntax and complex UDC codes

Reality check

Complex subjects, complex notations

```
−<rdf:RDF>
  −<rdf:Description>
    −<dc:title>
        Ageism in the workplace: stereotype threat, work disengagement and organizational disidentification
        among older workers
      </dc:title>
      <dc:creator>Oliveira, Eduardo André da Silva</dc:creator>
      <dc:subject xsi:type="dcterms:UDC">005.95/.96(043), por</dc:subject>
      <dc:subject xsi:type="dcterms:UDC">331.101-053.9, por</dc:subject>
      <dc:subject xsi:type="dcterms:UDC">316.346-053.9(043), por</dc:subject>
      <dc:subject xsi:type="dcterms:UDC">316.47(043), por</dc:subject>
      <dc:description>Monografia</dc:description>
    −<dc:description>
        Tese dout. Business and Management Studies Concentration (Organisation and Human Resources),
        Faculdade de Economia, Universidade do Porto, 2016
      </dc:description>
      <dc:publisher>[s.n.]</dc:publisher>
      <dc:publisher>Porto</dc:publisher>
```

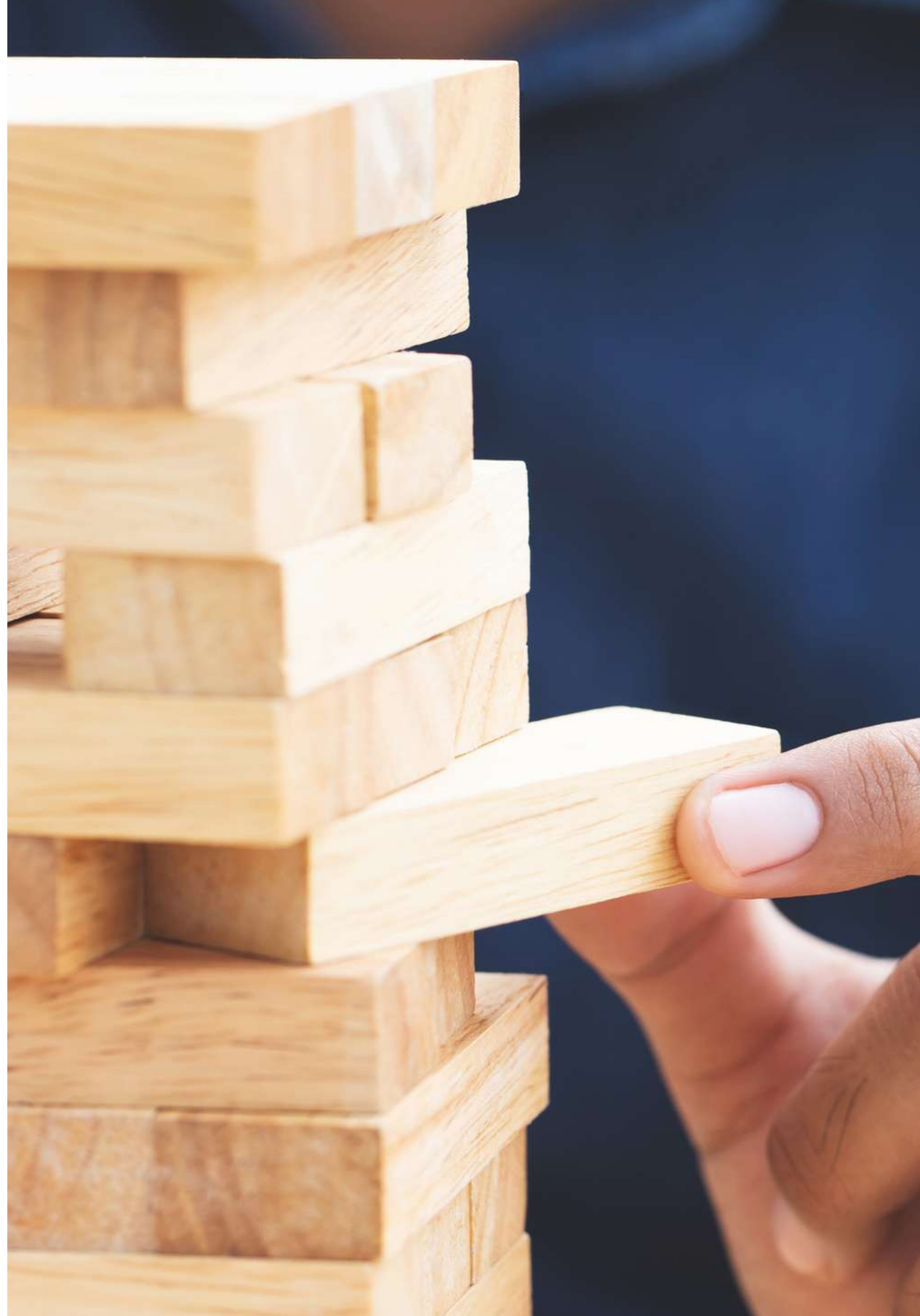
Source: https://opendata.bnportugal.gov.pt/



Reality check

Complex subjects, complex notations

005.95/.96	HR management	main classes
331.101	Labour	
316.47	Social relations	
-053.9	Senior citizens	common auxiliaries
(043)	Theses. Dissertations	



Reality check

Not all data can be linked

1. cancelled notations (concepts moved from one place to another)
2. complex syntax and complex UDC codes
3. licence paywall (not all of UDC data can be published as LOD)

Reality check

Licence issue: LD vs LOD

UDC Linked open data (LOD)

UDC Summary (57 languages)

CC BY-NC-ND 4.0

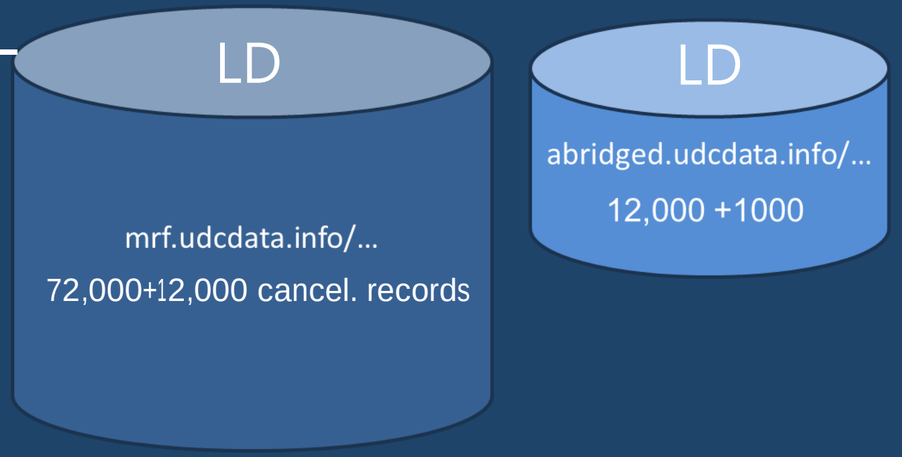
UDC RDF stores

UDC Linked data (LD) behind paywall (licensees only)

UDC Master Reference File (MRF)

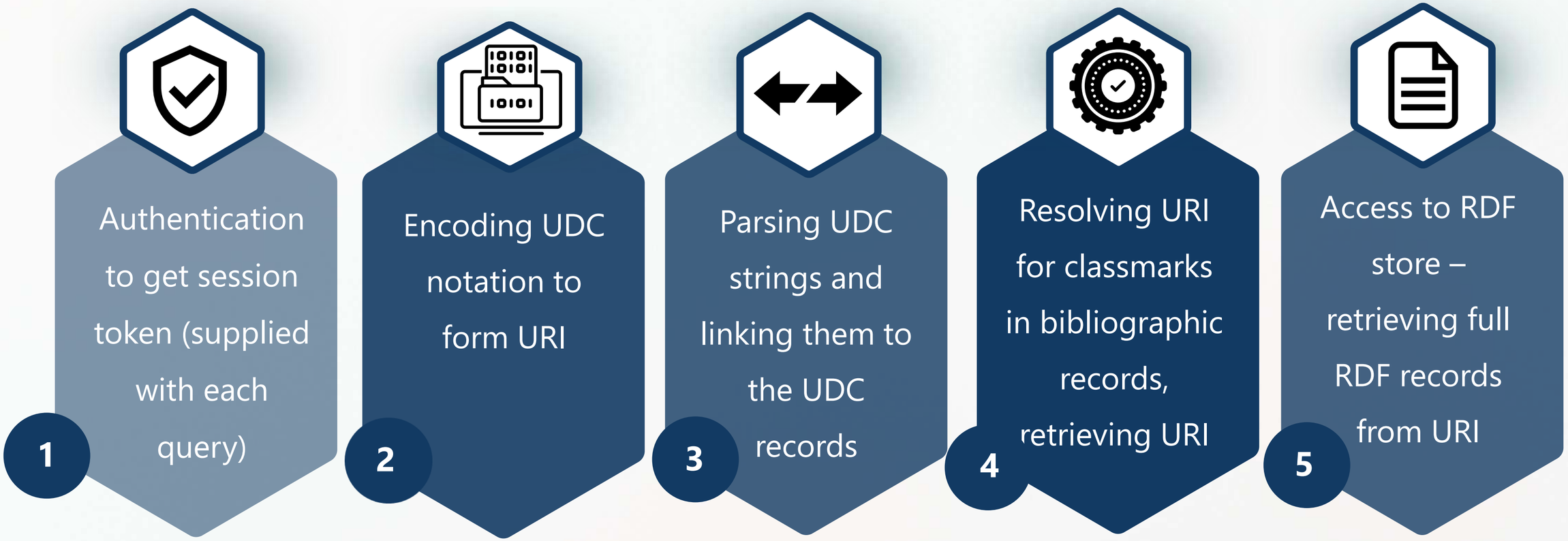
UDC Abridged

Further complexity with UDC MRF in different languages – shared ownership



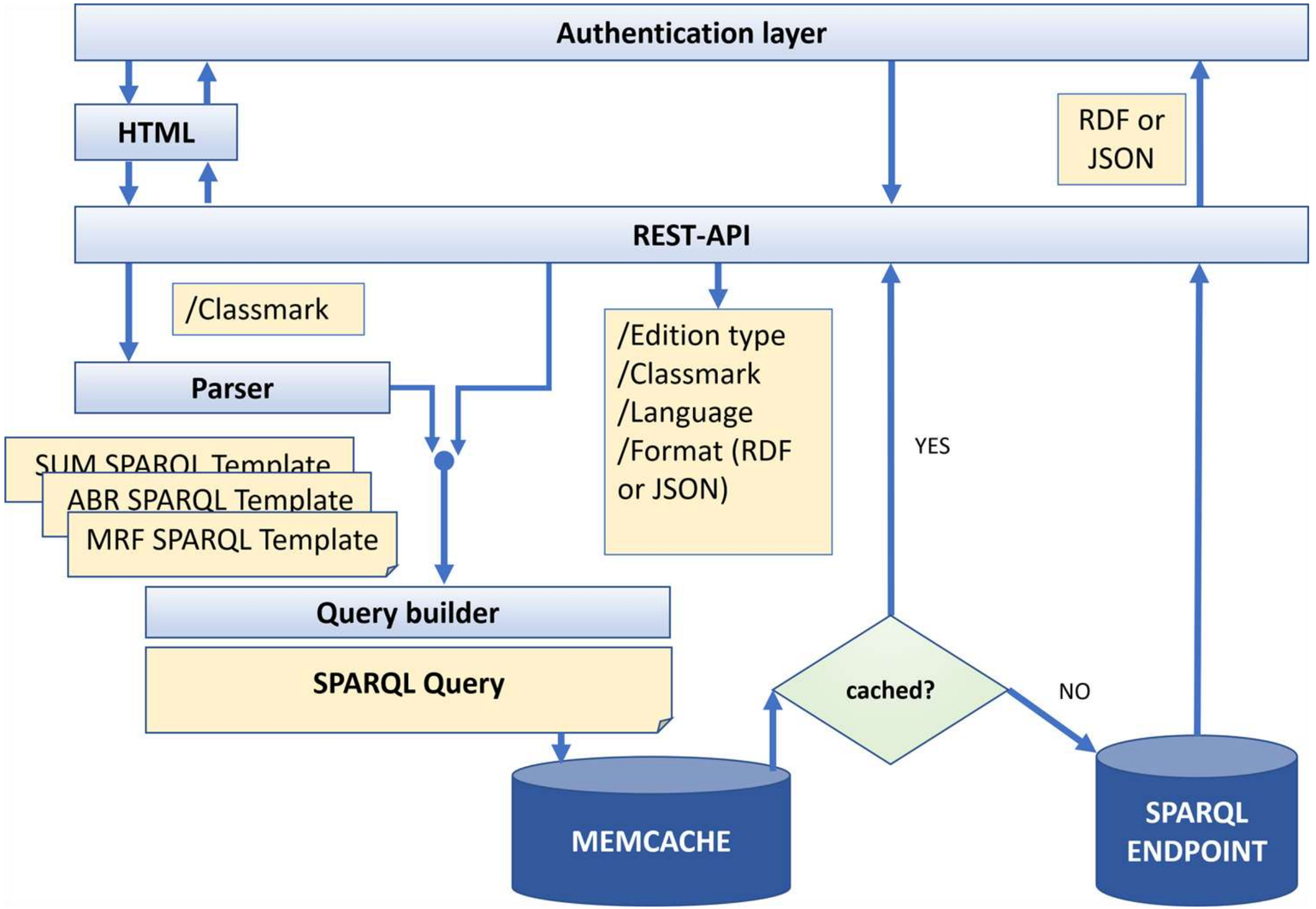
Solution: UDC look-up service

Three RDF stores containing all active and cancelled records with restricted SPARQL/Web API access



… if library classmark retrieves cancelled records, redirection is provided for the current concept placement

UDC look-up service architecture



UDC look-up service: specs and documents

URI format:

udcdata.info/MRF93/(492)

Encoded as: udcdata.info/MRF93/_or_492_cr_

Schema(s):

RDFS (RDF Schema)

SKOS (extended to accommodate historical data, additional 3 data elements)

Parsing:

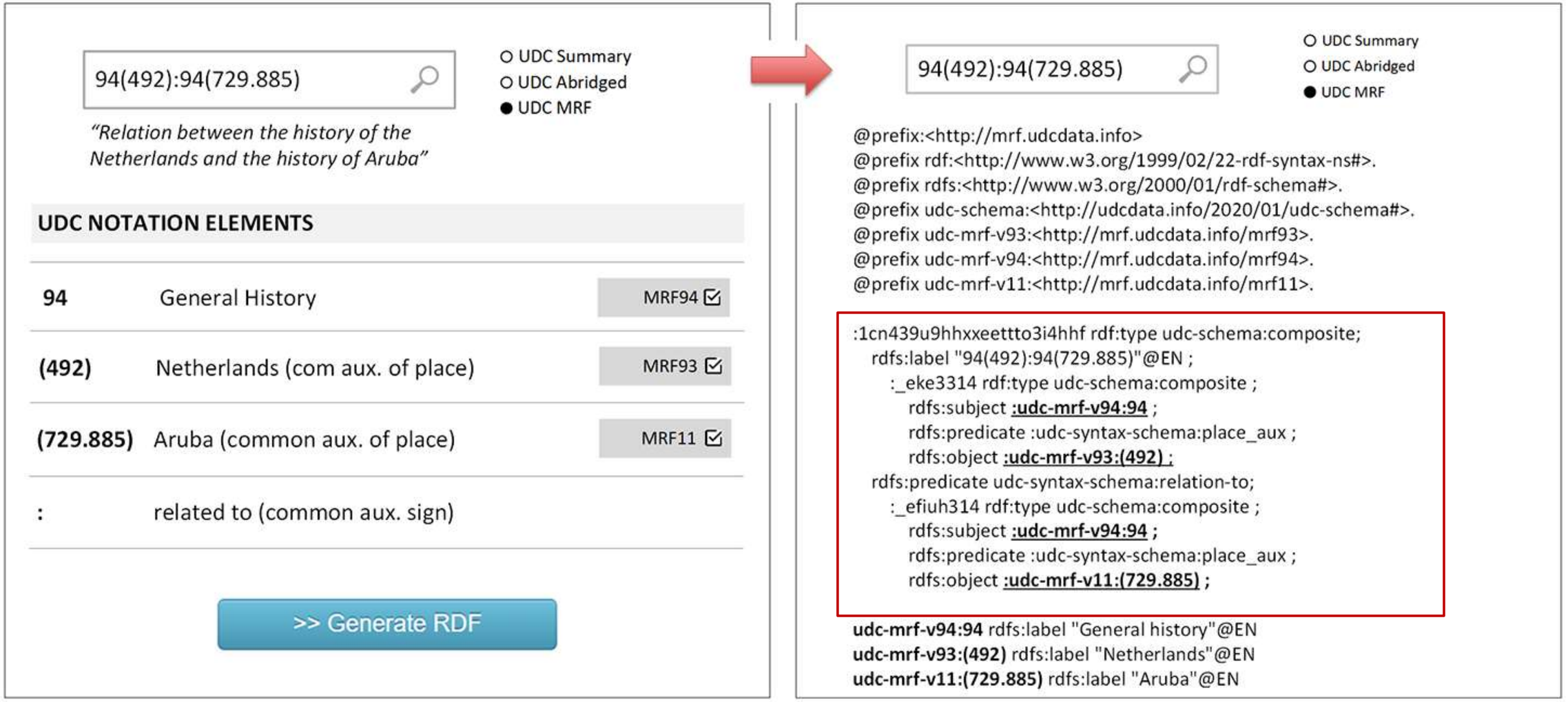
UDC notation interpreter, automata based on a series of algorithms that splits UDC notation

Dataset access policy (licence paywall):

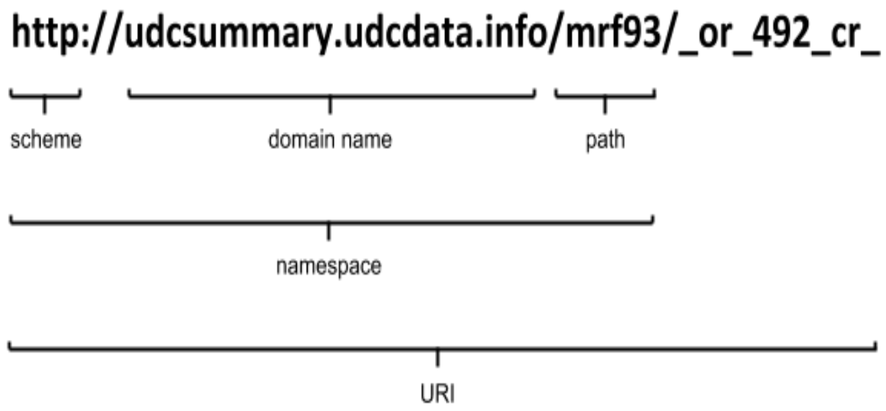
Authentication process is handled by standard shared and private authentication keys (tokens)

UDC number (notation)	skos:notation
class identifier (URI)	skos:Concept
broader class (URI)	skos:broader
caption	skos:prefLabel
including note	skos:note
application note	skos:note
scope note	skos:scopeNote
examples	skos:example
see also reference	skos:related
	skos : historyNote
dcterms: created	+UDC SUBPROPERTIES
dcterms: deprecated	ReplacedBy
dcterms: modified	NotationPreviouslyUsedFor
	ConceptPreviouslyAt

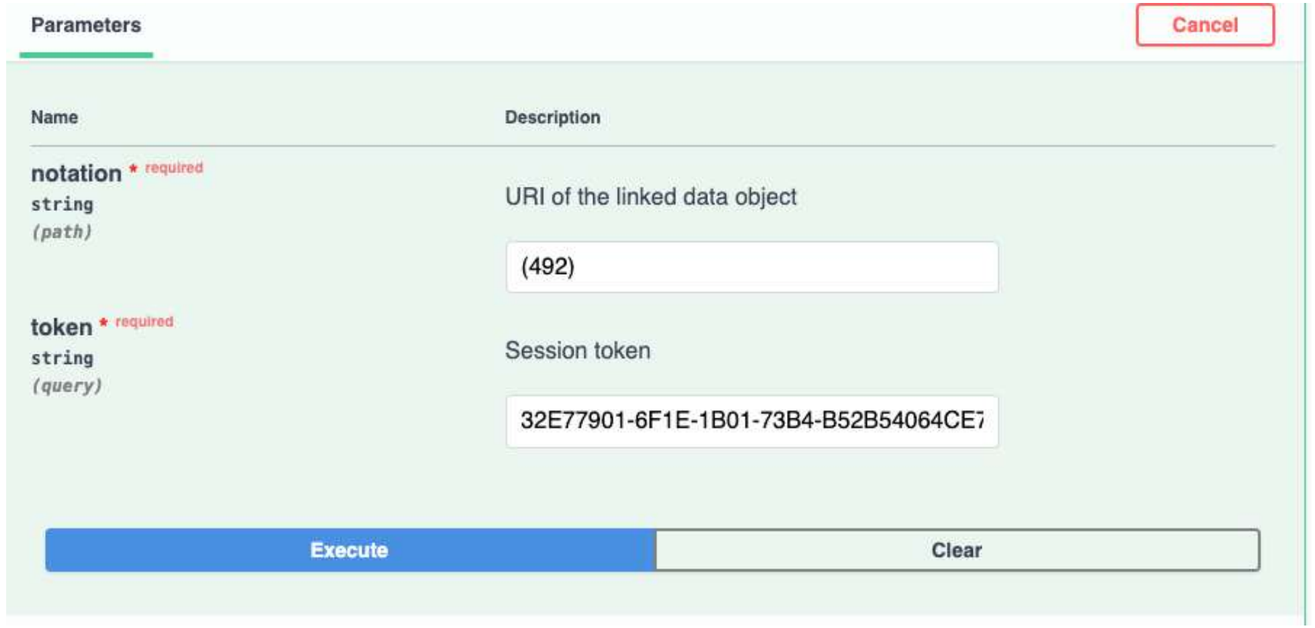
UDC look-up service: parsing scenario

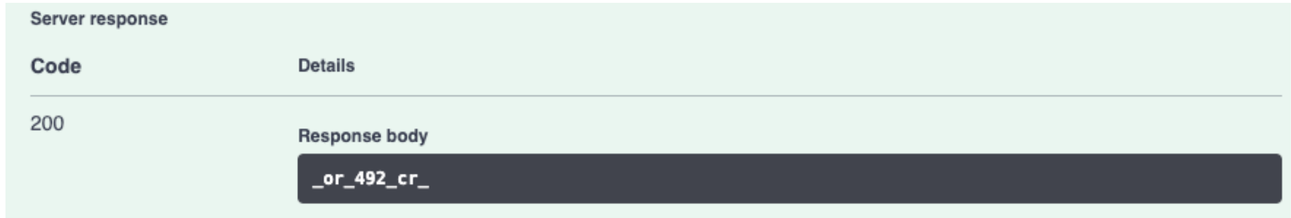


UDC look-up service: classmark (492) in URI encoding



UDC notation encoding API





Testing: Query the URI for notation 94(100)

POST `/api/lookup/{notation}` Query the active and cancelled URIs for the supplied UDC notation.

Query the active and cancelled URIs for the supplied UDC notation.

Parameters

Name	Description
notation * required string *(path)*	UDC notation 94_or_100_cr_
token * required string *(query)*	Session token 2C0E8BCA-C98B-175F-FA4E-6052E556547

Execute

Code	Details
200	**Response body** `{"uris": [{"uri": "http:\/\/udcdata.info\/mrf08\/94_or_100_cr_","active":true}]}` **Response headers**

Testing: Returning RDF record

Server response

Code	Details
200	Response body

```
<?xml version="1.0"?>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:udc="http://udcdata.info/udc-schema#">
    <skos:Concept rdf:about="http://www.udcdata.info/mrf08/94_or_100_cr_">
      <skos:inScheme rdf:resource="http://udcdata.info/udc-schema"/>
      <skos:broader rdf:resource="http://udcdata.info/mrf94/94"/>
      <skos:notation>94(100)</skos:notation>
      <skos:prefLabel xml:lang="en">World history (chronological summation of facts)</skos:prefLabel>
      <skos:example xml:lang="en">94(100)&quot;.../05&quot; Ancient history in general. History of ancient people</skos:example>
      <skos:example xml:lang="en">94(100)&quot;05/...&quot; Mediaeval and modern history in general</skos:example>
      <skos:example xml:lang="en">94(100)&quot;1914/1918&quot; First World War, 1914-18 (as world history)</skos:example>
      <skos:example xml:lang="en">94(100)&quot;1939/1945&quot; Second World War, 1939-45 (as world history)</skos:example>
    </skos:Concept>
  </rdf:RDF>
```

UDC look-up service

Where are we now?

- All 20 versions of UDC MRF released since 1993 held in an RDF store (Virtuoso)

  Now adding 3,000 old UDC records from the 1980s

- Currently
  - testing authentication layer
  - expanding parsing algorithms to include new UDC syntax symbols (to be introduced in the next release of MRF13 (by the end of 2025)
  - designing HTML interface (for humans)

- UDC Summary

  RDF store (LOD) is being updated with new data and cancellations

  (release planned by July 2025)

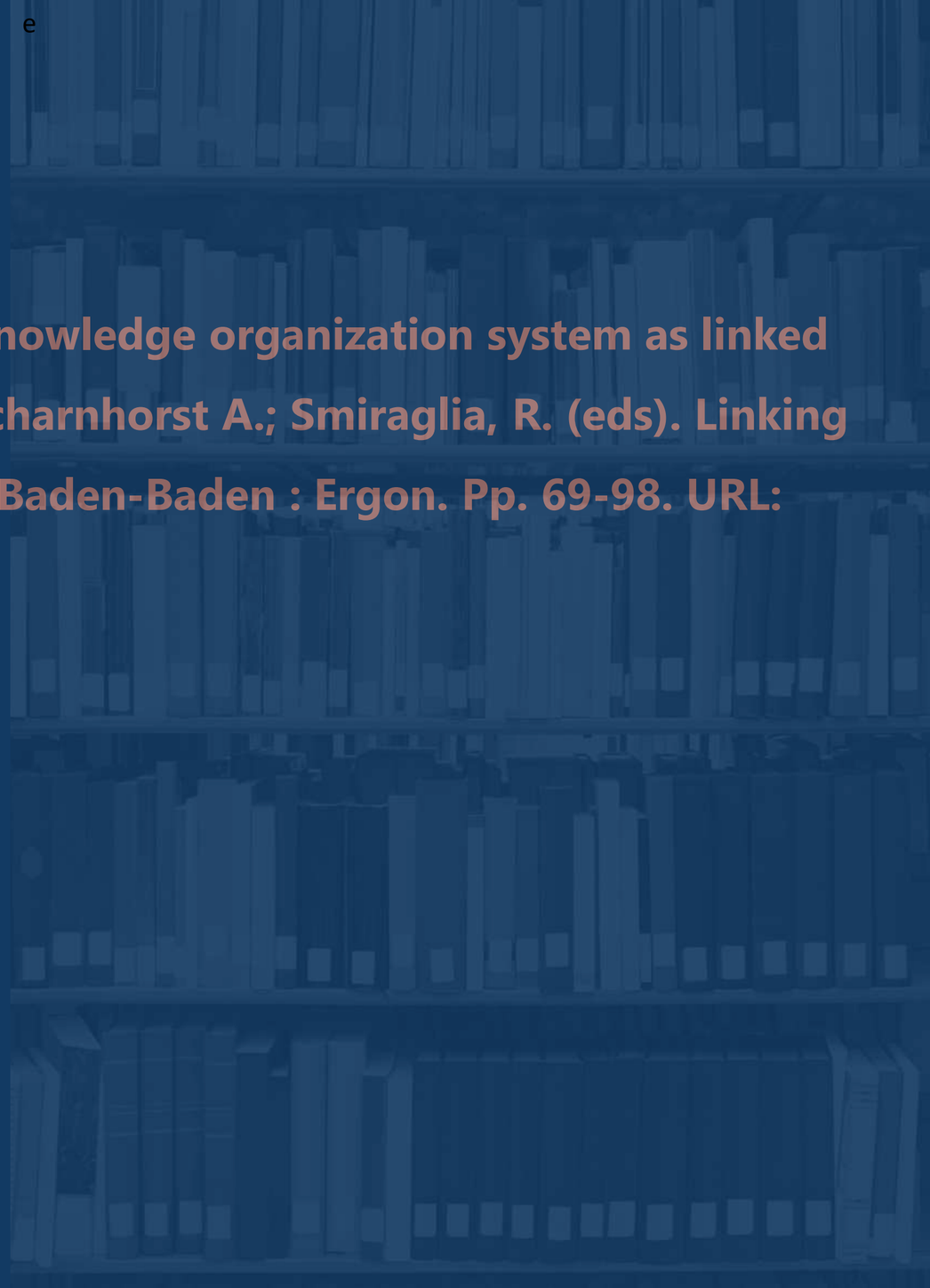
In conclusion

Challenges	Solutions
UDC longevity and change	Inclusion of historical UDC data and concordances between old and new classes; UDC version-based URI
Structural complexity	Provision for expressing UDC syntax in RDF and the use blank nodes to allow grouping of notational elements
Large user base & amount of legacy data	Provision for parsing, resolving and identifying UDC notations within a look-up service
Data ownership	Provision for managing both LOD data set and LD datasets behind the paywall (authentication)

e

More info

Slavic, A.; Siebes, R.; Scharnhorst, A. (2021). Publishing a knowledge organization system as linked data: the case of the Universal Decimal Classification. In: Scharnhorst A.; Smiraglia, R. (eds). Linking knowledge: linked open data for knowledge organization. Baden-Baden : Ergon. Pp. 69-98. URL: https://arxiv.org/abs/2205.01395



Thank you

aida.slavic@udcc.org
ana.vukadin@udcc.org