KMSKA, artinflanders.be, Dominique Provost, public domain

The role of E-ARK in meemoo's ingest

10 years of E(-)ARKiving

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Today’s (story) arc

Who is meemoo?

A bit of context

Our search for a SIP

Introducing the ‘meemoo SIP’

Ingest SIPs into our archive

Lessons learned & way forward

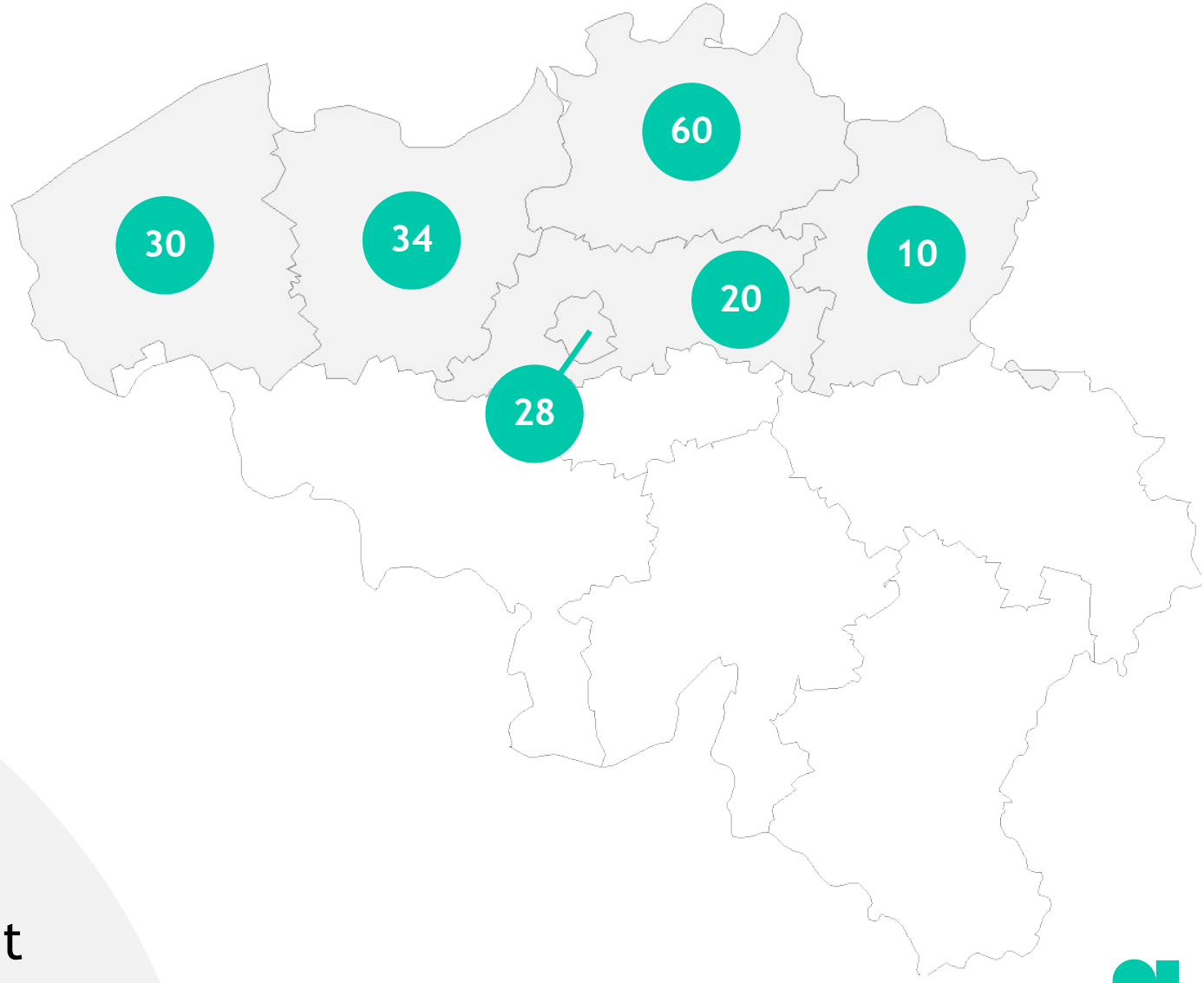
meemoo’s service provision

- **Digitisation**, **digital archiving** and **management** of archival materials
- Make content **accessible and usable**
- Actively **gather and share expertise** on digital archive operations
- **Advise** on digital heritage processes

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

These figures are from 31 December 2023.

Collaboration



182 content partners (CP) in culture, media and government

meemoo
FLEMISH INSTITUTE FOR ARCHIVES

These figures are from 31 December 2023.

Content partners in different sectors

54 performing arts	12 government institutions
50 museums	7 heritage libraries
24 archives	3 national broadcasters
20 heritage societies	2 sector institutes
10 regional broadcasters	

meemoo
FLEMISH INSTITUTE FOR ARCHIVES

In figures

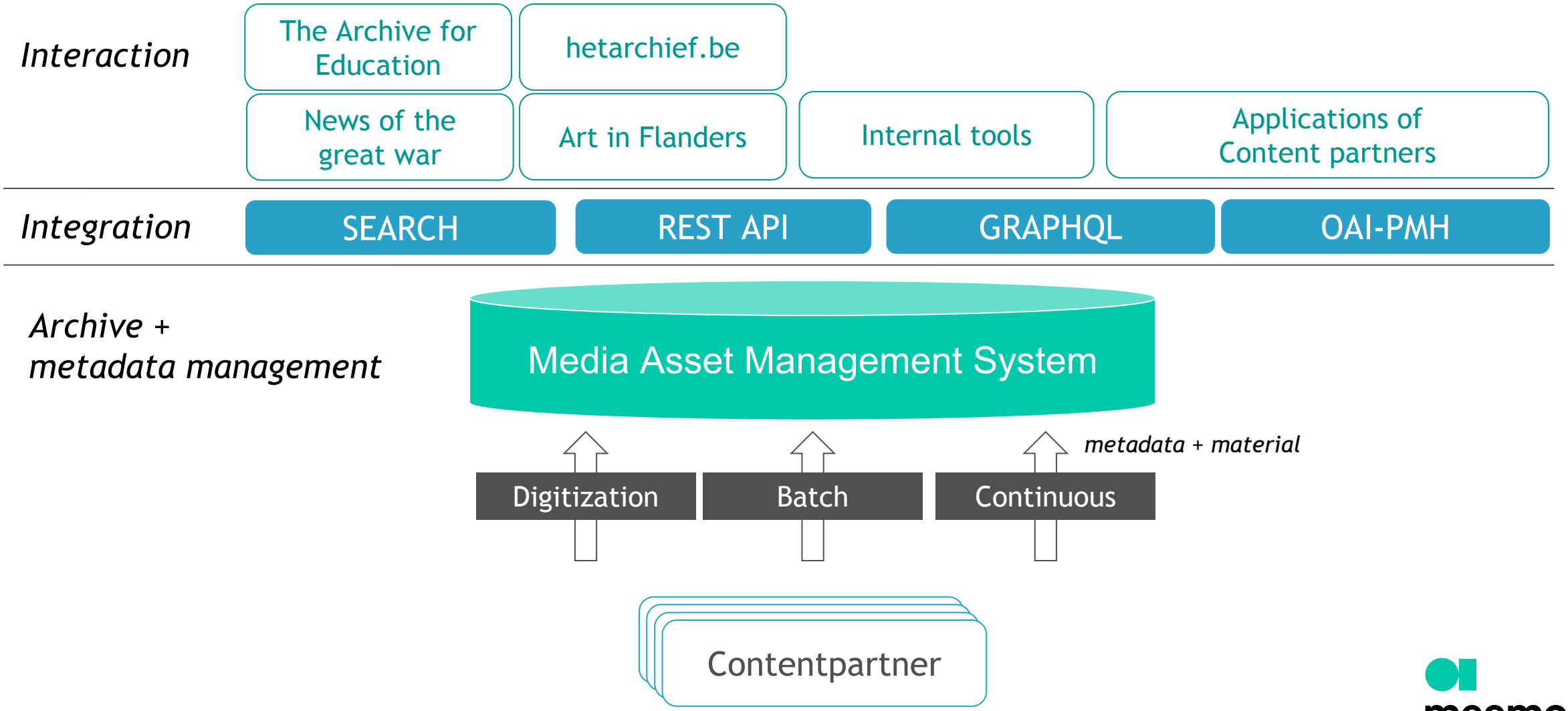
> 8 million
objects in our archive system

> 540,000
audiovisual carriers transferred to our archive system

nearly 185,000
user accounts at The Archive for Education at end of 2022-2023 academic year

All these figures except for education are from 31 December 2023.

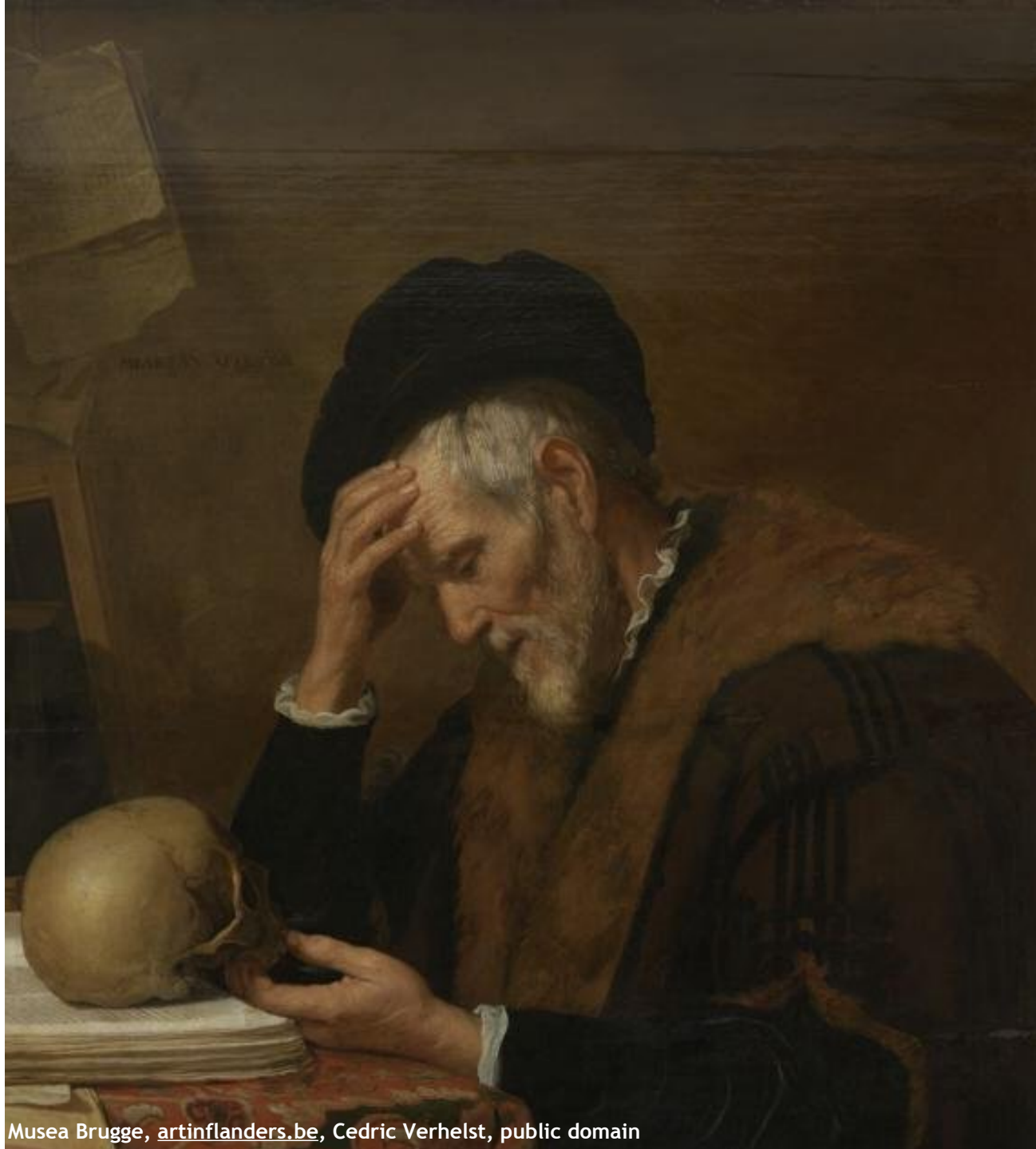
meemoo
FLEMISH INSTITUTE FOR ARCHIVES



meemoo
FLEMISH INSTITUTE FOR ARCHIVES

A bit of context

how we came to the realization that we needed a SIP



Musea Brugge, artinflanders.be, Cedric Verhelst, public domain

First SIP practice in Digitisation

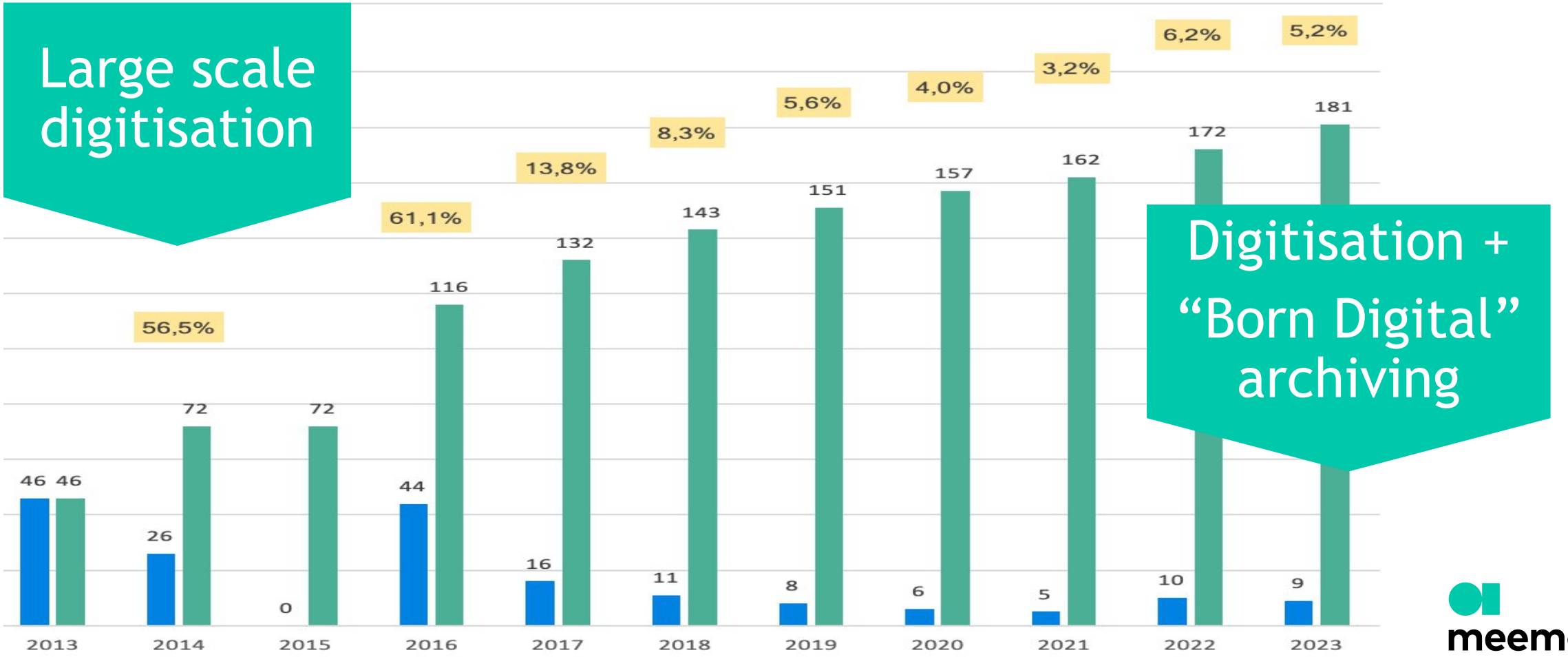
Digitisation via **external digitization companies**

Manageable & simple **packaging rules**

Vendor-imposed SIP (with a descriptive metadata twist)

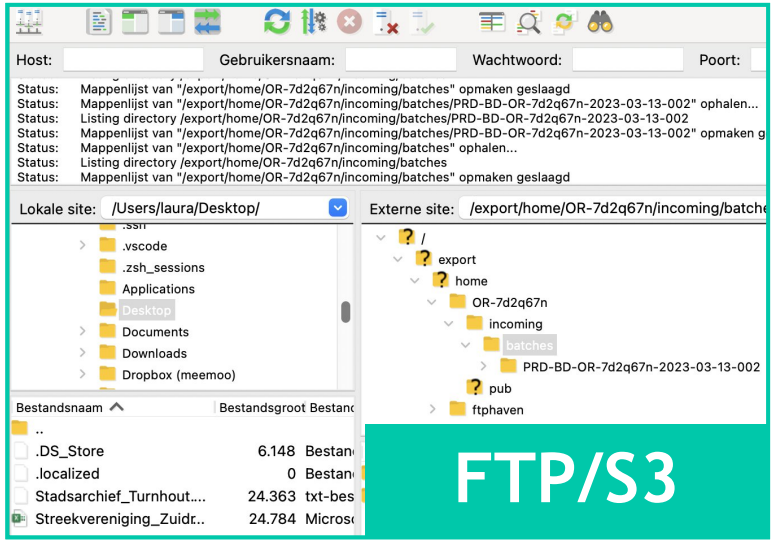
meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

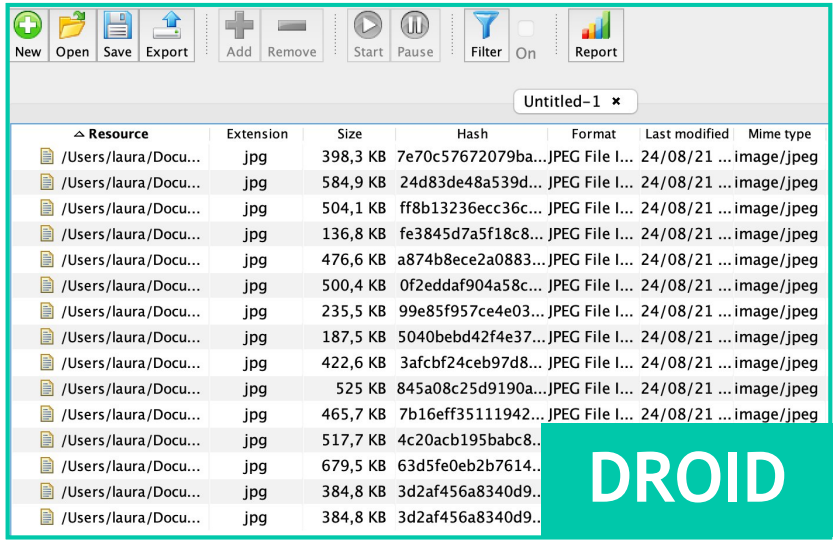
Rise in number of partners



meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Influx of digital collections

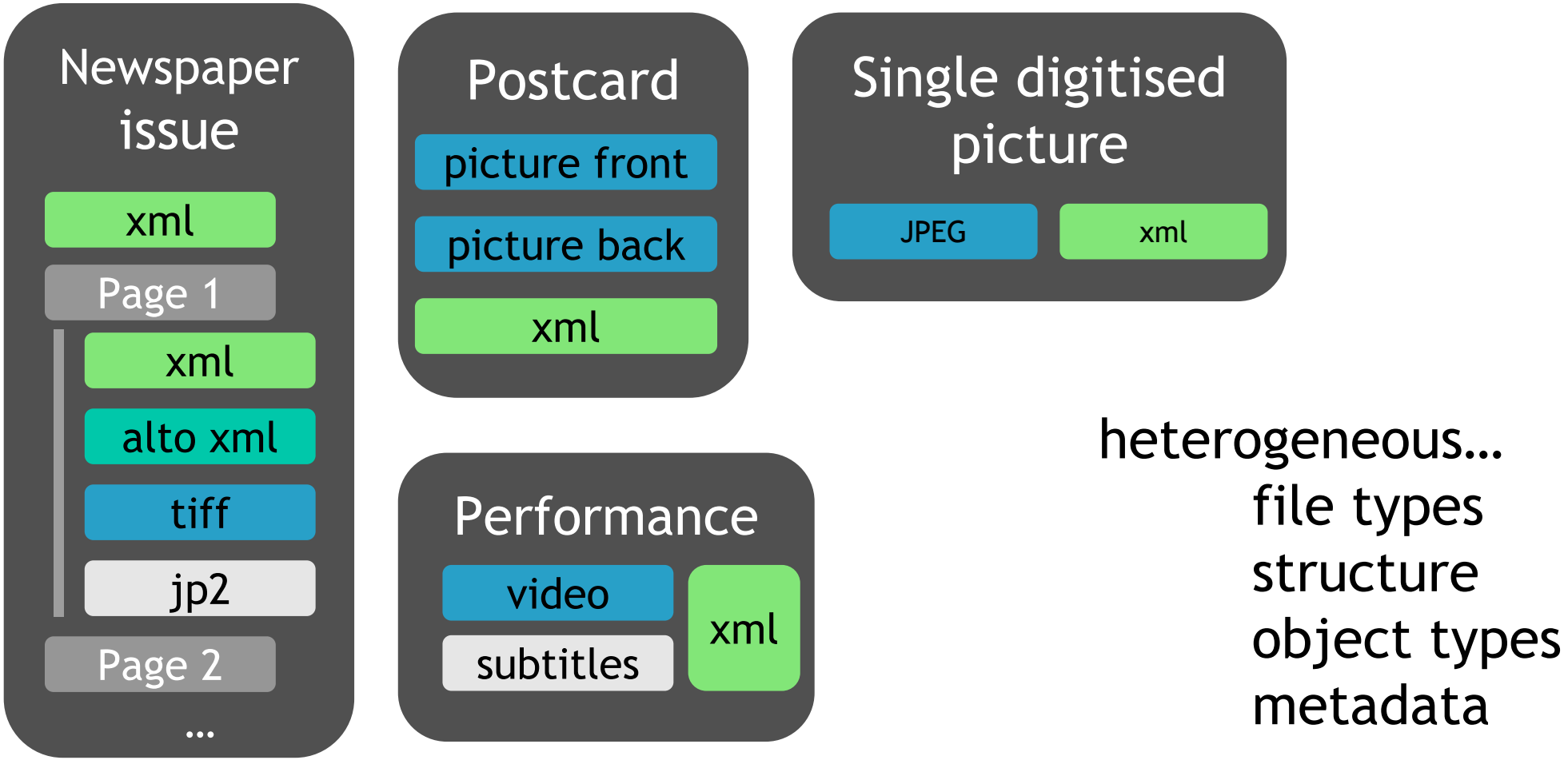




dc_identifier_localid	dc_title	dc_description	dcterms_created	dc_identifier_localids_Bestandsnaam	dc_titles_reeks	dc_titles_seizoen	dcterms_issued	dc_creators_Acteur
Hoofd lokale CP ID = bestandsnaam zonder extensie	*Hoofdtitel*	*Hoofdbeschrijving*	*Creatiedatum van het object, EDTF level 0 of 1. Zie developer.meemoo.be (datatypes)*	*Originele bestandsnaam van het object = hoofd lokale CP ID + extensie*	*Naam van de voorstelling of productie*	*Het seizoen van de voorstelling*	*Datum waarop het item is uitgegeven of uitgezonden. EDTF level 0 of 1. Zie developer.meemoo.be (datatypes)*	*De **acteur** in het object in kwestie*
Offtherecord_CP_1	Off the record	Vier jongeren tussen 14 en 18 jaar tonen zich zoals ze zijn. Fel, kwetsbaar, schaamteloos en vol lef. Het podium is een vrijplaats waar ze letterlijk en figuurlijk spelen met de vele versies van zichzelf. Als hoofdrolspelers in hun eigen documentaire. Ze zijn de stemmen van het heden, de volwassenen van de toekomst. Off the record is een thea touwtjes zelf in handen r zien en hoe ze denken (ni	2022-02-28	Offtherecord_CP_1.jpg	Off the record	2021 - 2022;2022 - 2023		Ben Herlaar Masina;Ziba Jafari;Mauro Michielsen;Mette Plysier

Descriptive metadata templates

More complex object structures



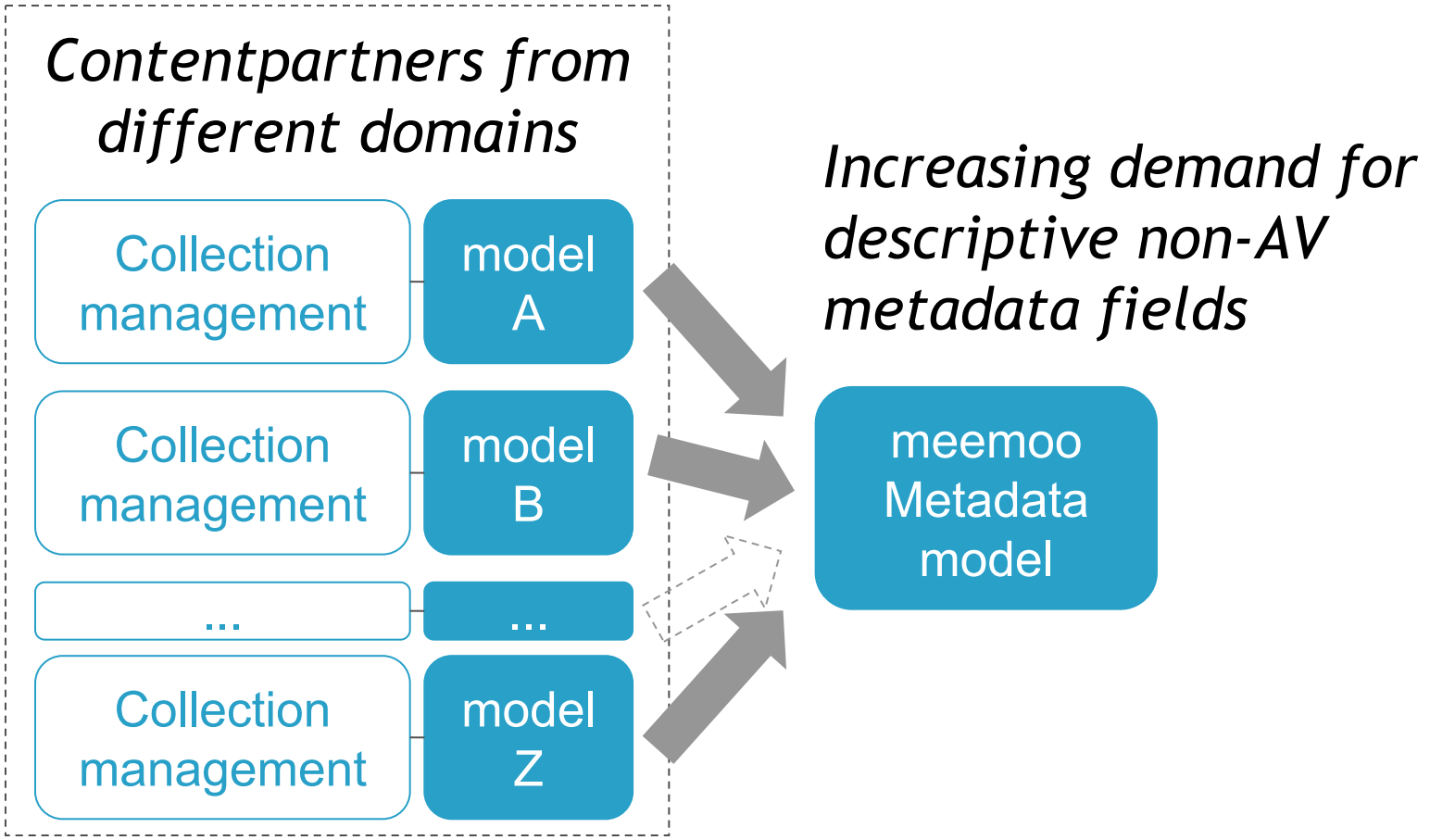
heterogeneous...

- file types
- structure
- object types
- metadata

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

More heterogeneity in metadata

Diverse metadata intake



meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Mitigation strategies (sort of)

- Divide digital influx into different workflows
- New vendor-specific SIPs for e.g. newspapers
- Providing custom services via project manager

→ All of this works (kind of), but increases complexity and overhead

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Setting sail for a (new) SIP

how we came aboard the (E-)ARK

Musea Brugge, artinflanders.be, Dominique Provost, public domain

Requirements

A new uniform SIP design should …

have a **formal and mature specification**

be embedded in **standards**

be technologically aligned with the **media and cultural heritage** sector

be **flexible and extensible**

Options we explored

	Standards	Specification	Mature	Verbosity	Community support	Sector alignment
RO-Crate						
PCDM						
BagIt						
E-ARK						

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Closing the E-ARK and meemoo gap

Flexibility can lead to complexity, esp. with not very tech savvy partners

Need for more rules wrt. to object structure, but mostly descriptive metadata

Translate E-ARK spec into our own meemoo SIP specification

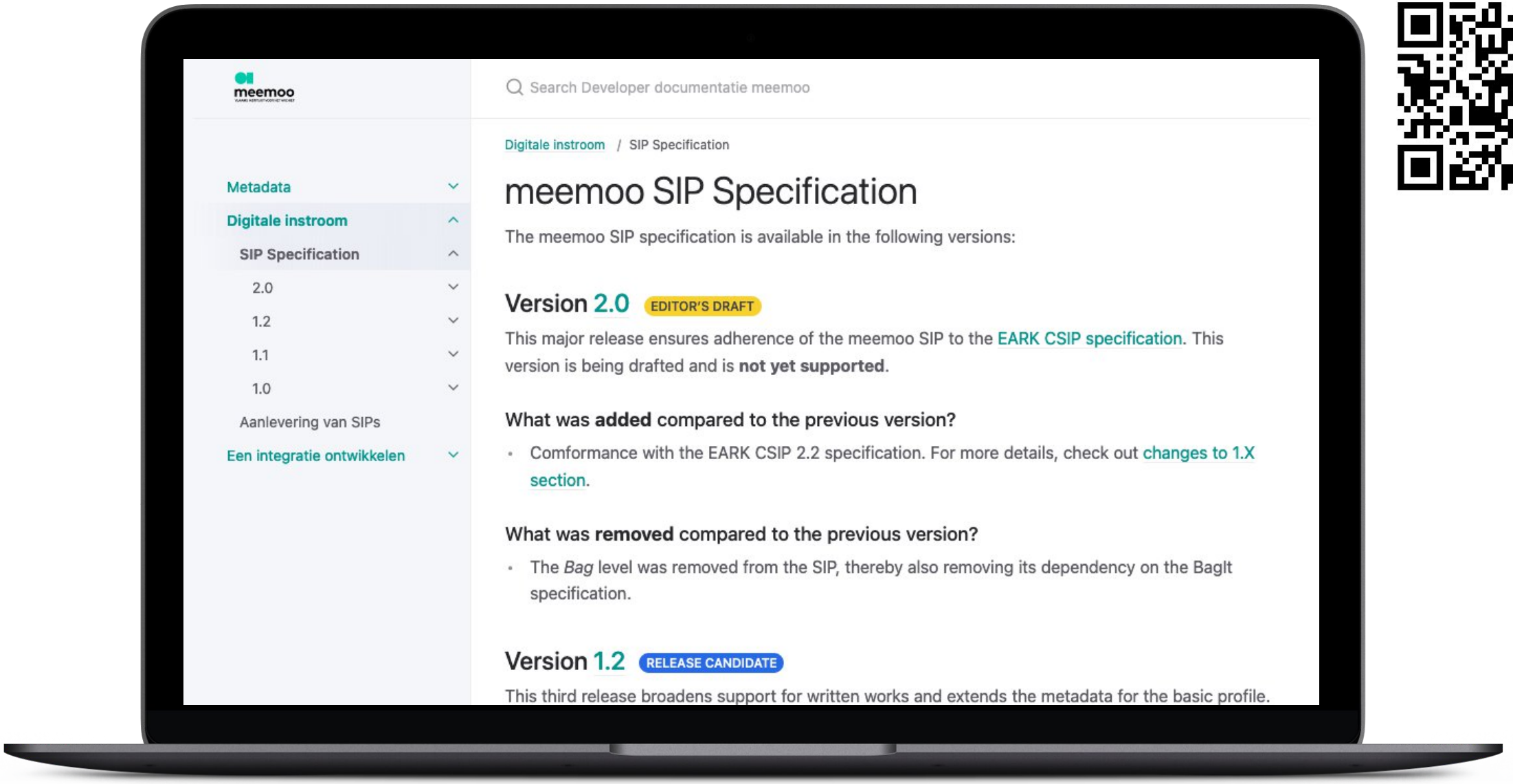
- `terminology` + `core concepts` sections
- narrowly following the hierarchical structure
- conceived the concept of ‘content profiles’: additional use-case driven constraints, often domain or material specific

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

A (deep) dive into the ‘meemoo SIP’

how we put it in a Bag only to remove it afterwards

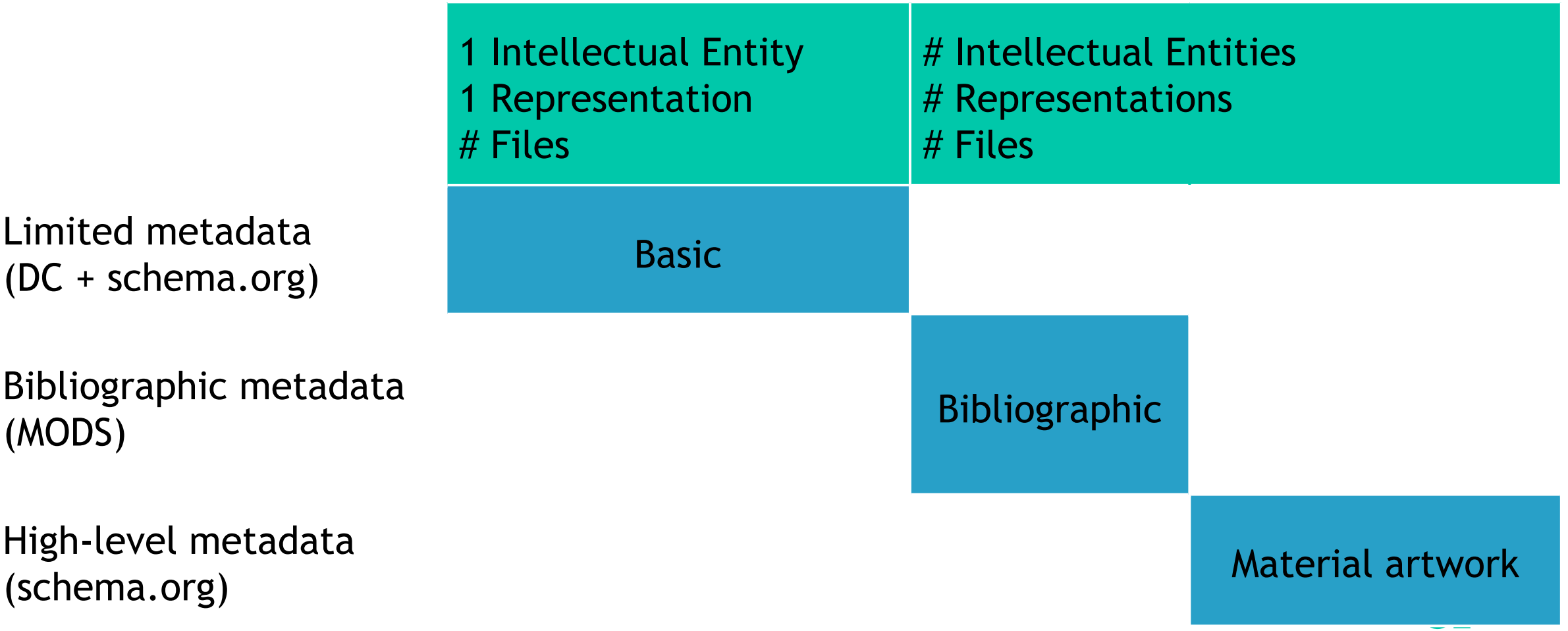
Unknown artist, The Illustrated London News, public domain (via Wikimedia Commons)

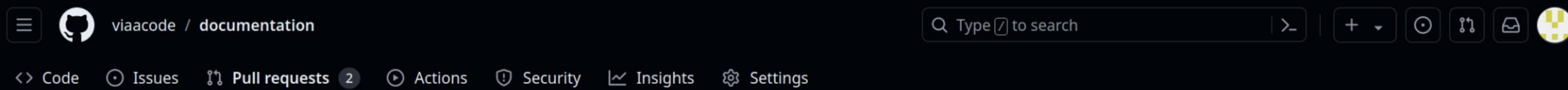


https://data.hetarchief.be/id/sip/

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Content profiles





Filters is:pr is:closed SIP

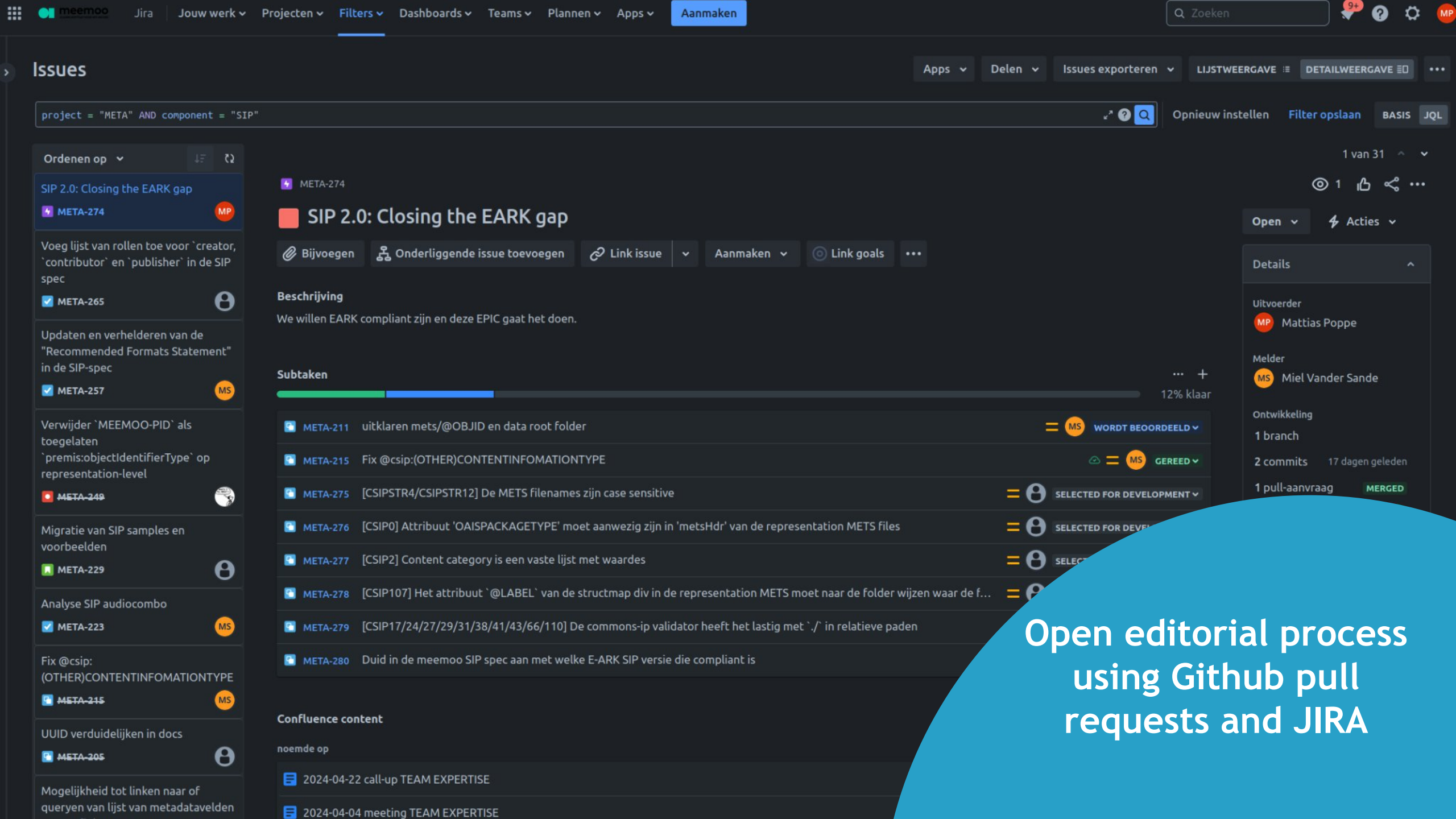
Labels 9 Milestones 0 New pull request

Clear current search query, filters, and sorts

1 Open 26 Closed Author Label Projects Milestones Reviews Assignee Sort

- Meta 274 sip 2 0 closing the eark gap — 55
  #130 by mielvds was merged 2 weeks ago
- META-266 - Fix (C)SIP namespace — 25
  #125 by spacid was merged on Mar 4 • Approved
- DEV-2286: External documentation for S3 — 12
  #119 by maartends was merged on Jan 29 • Approved
- Meta 171 sip versie 1 2
  #110 by mielvds was merged on Nov 2, 2023
- Meta 230 aanvullen sip voor metadata give 3 d
  #100 by mielvds was merged on Sep 20, 2023
- SIP spec: art form en art medium moeten van type lang string zijn en niet string
  #91 by mielvds was merged on Jun 13, 2023 • Approved
- META-204 Remake bags
  #72 by spacid was merged on Apr 6, 2023 • Approved
- Add unique language contraint to spec
  #65 by mielvds was merged on Mar 27, 2023 • Approved
- META-153 Fixes contractions obligation and cardinality Representation PREMIS

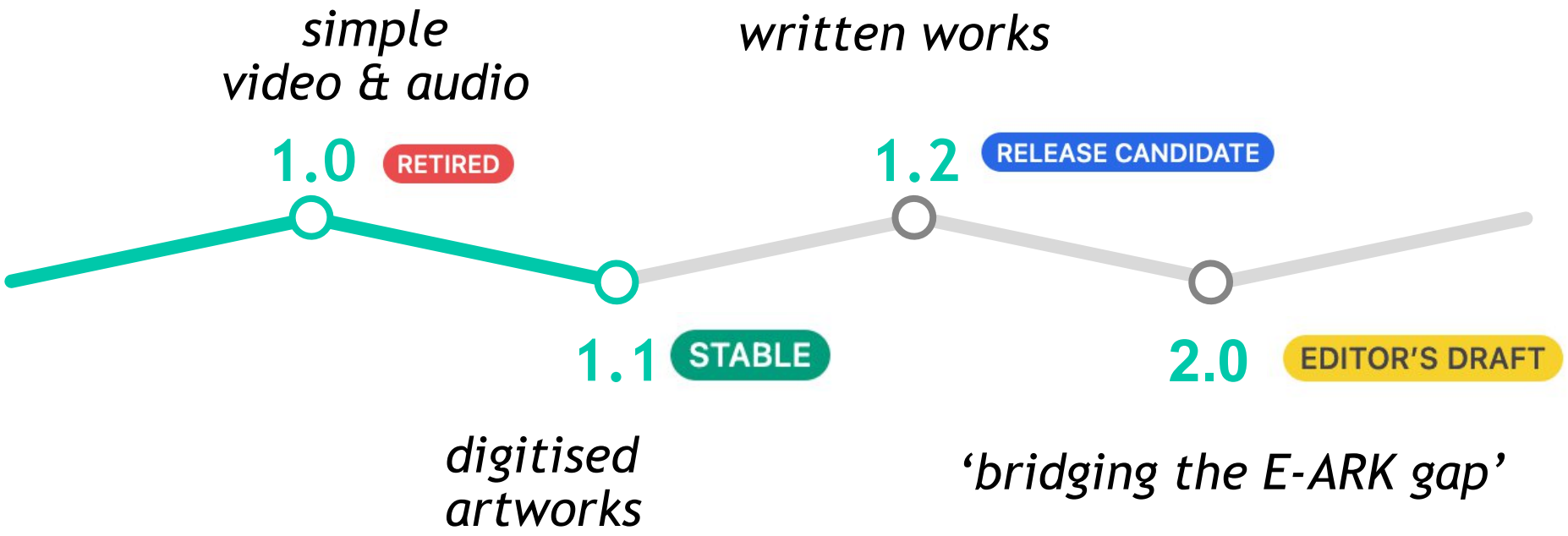
Open editorial process using Github pull requests and JIRA


Jira
Jouw werk
Projecten
Filters
Dashboards
Teams
Plannen
Apps
Aanmaken
Zoeken
Issues
Apps
Delen
Issues exporteren
LIJSTWEERGAVE
DETAILWEERGAVE
project = "META" AND component = "SIP"
Opnieuw instellen
Filter opslaan
BASIS
JQL
Ordenen op
SIP 2.0: Closing the EARK gap
META-274
Voeg lijst van rollen toe voor `creator`, `contributor` en `publisher` in de SIP spec
META-265
Updaten en verhelderen van de "Recommended Formats Statement" in de SIP-spec
META-257
Verwijder `MEEMOO-PID` als toegelaten `premis:objectIdentifierType` op representation-level
META-249
Migratie van SIP samples en voorbeelden
META-229
Analyse SIP audiocombo
META-223
Fix @csip:(OTHER)CONTENTINFOMATIONTYPE
META-215
UUID verduidelijken in docs
META-205
Mogelijkheid tot linken naar of queryen van lijst van metadatavelden
1 van 31
META-274
SIP 2.0: Closing the EARK gap
Bijvoegen
Onderliggende issue toevoegen
Link issue
Aanmaken
Link goals
Beschrijving
We willen EARK compliant zijn en deze EPIC gaat het doen.
Subtaken
12% klaar
META-211 uitklaren mets/@OBJID en data root folder WORDT BEOORDEELD
META-215 Fix @csip:(OTHER)CONTENTINFOMATIONTYPE GEREED
META-275 [CSIPSTR4/CSIPSTR12] De METS filenames zijn case sensitive SELECTED FOR DEVELOPMENT
META-276 [CSIP0] Attribuut 'OAISPACKAGETYPE' moet aanwezig zijn in 'metsHdr' van de representation METS files
META-277 [CSIP2] Content category is een vaste lijst met waardes
META-278 [CSIP107] Het attribuut `@LABEL` van de structmap div in de representation METS moet naar de folder wijzen waar de f...
META-279 [CSIP17/24/27/29/31/38/41/43/66/110] De commons-ip validator heeft het lastig met `./` in relatieve paden
META-280 Duid in de meemoo SIP spec aan met welke E-ARK SIP versie die compliant is
Confluence content
noemde op
2024-04-22 call-up TEAM EXPERTISE
2024-04-04 meeting TEAM EXPERTISE
Open
Acties
Details
Uitvoerder
Mattias Poppe
Melder
Miel Vander Sande
Ontwikkeling
1 branch
2 commits
17 dagen geleden
1 pull-aanvraag
MERGED
Open editorial process using Github pull requests and JIRA

Support level through publication statuses

EDITOR'S DRAFT	This version of the specification is still being developed.
RELEASE CANDIDATE	This version of the specification is ready for *public review* and *implementation*. The text can still undergo changes and meemoo does not give any guarantees that material can be supplied in this way.
STABLE	This version of the specification will no longer change and has been fully implemented. Meemoo guarantees that material supplied in this way can be archived.
DEPRECATED	This version of the specification is outdated and will no longer be supported in the future. Its use is not recommended by meemoo.
RETIRED	This version of the specification is no longer supported by meemoo. Material can no longer be supplied in this way.

Versioning



meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Move editing & versioning into governance

Release cycle to avoid ad-hoc alterations

Connect with partners, gather input & cover their needs

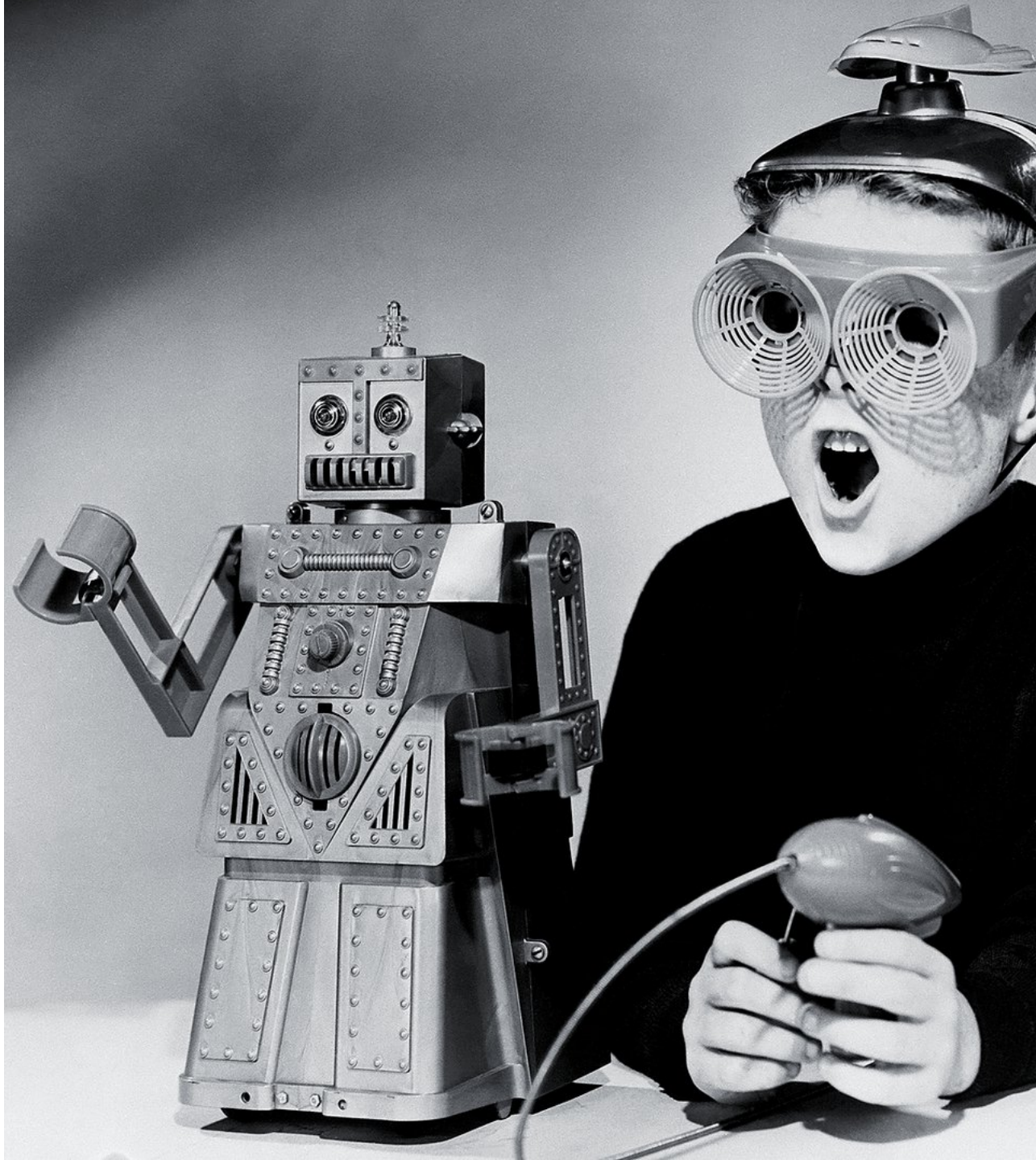
Predictable support for both sides

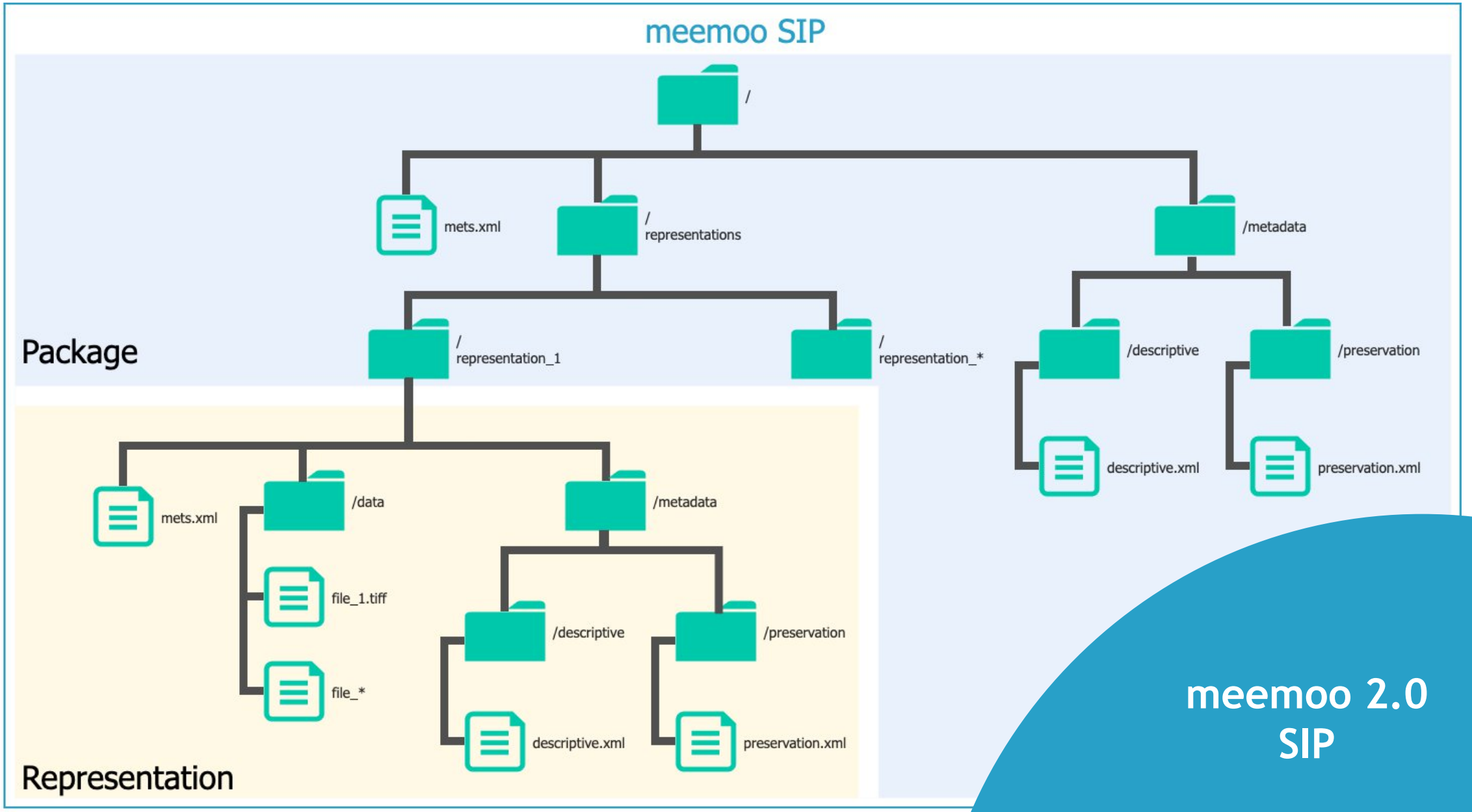
Still ongoing work

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Ingesting a ‘meemoo SIP’

how we transform it to a graph for semantic validation





meemoo 2.0 SIP

Ingesting a ‘meemoo SIP’

Validating the SIP as a graph

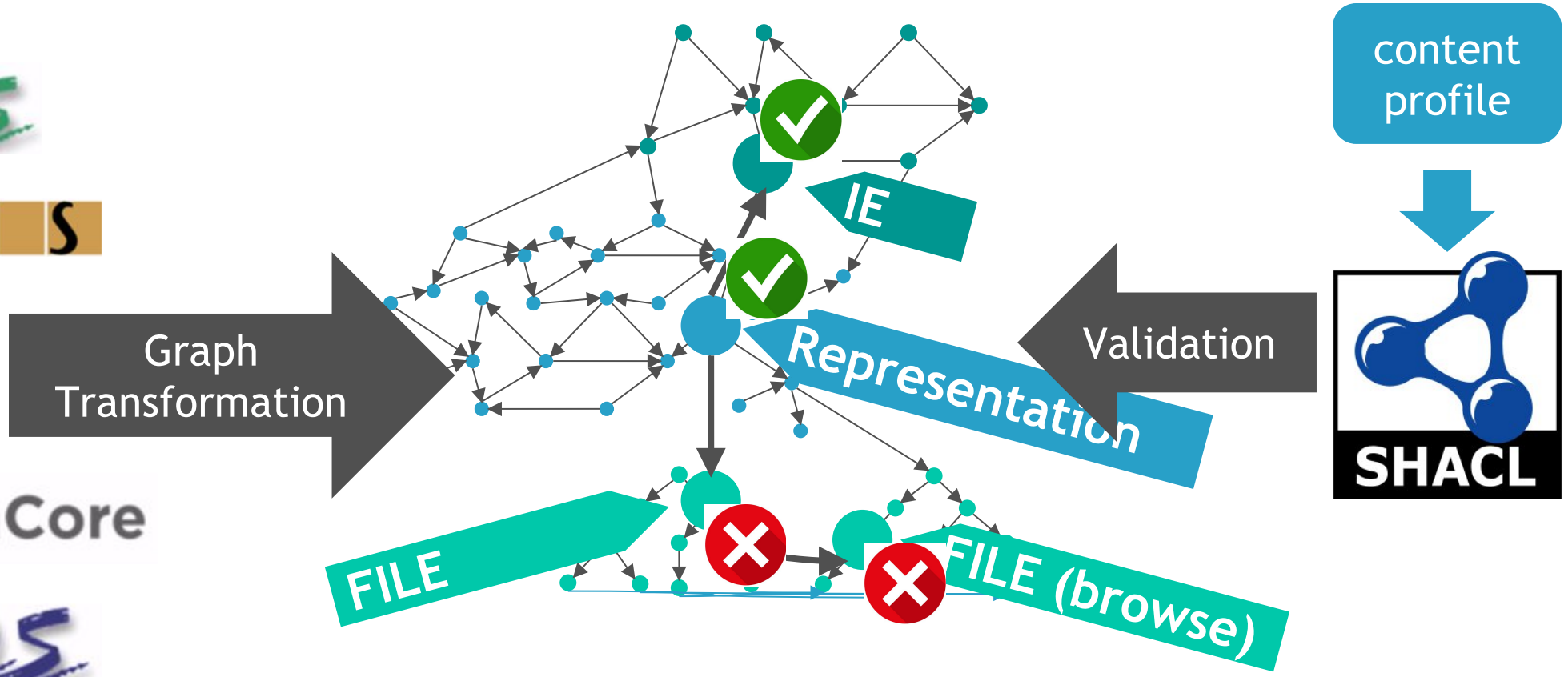
METS

PREMIS

DublinCore

MODS

Schema.org

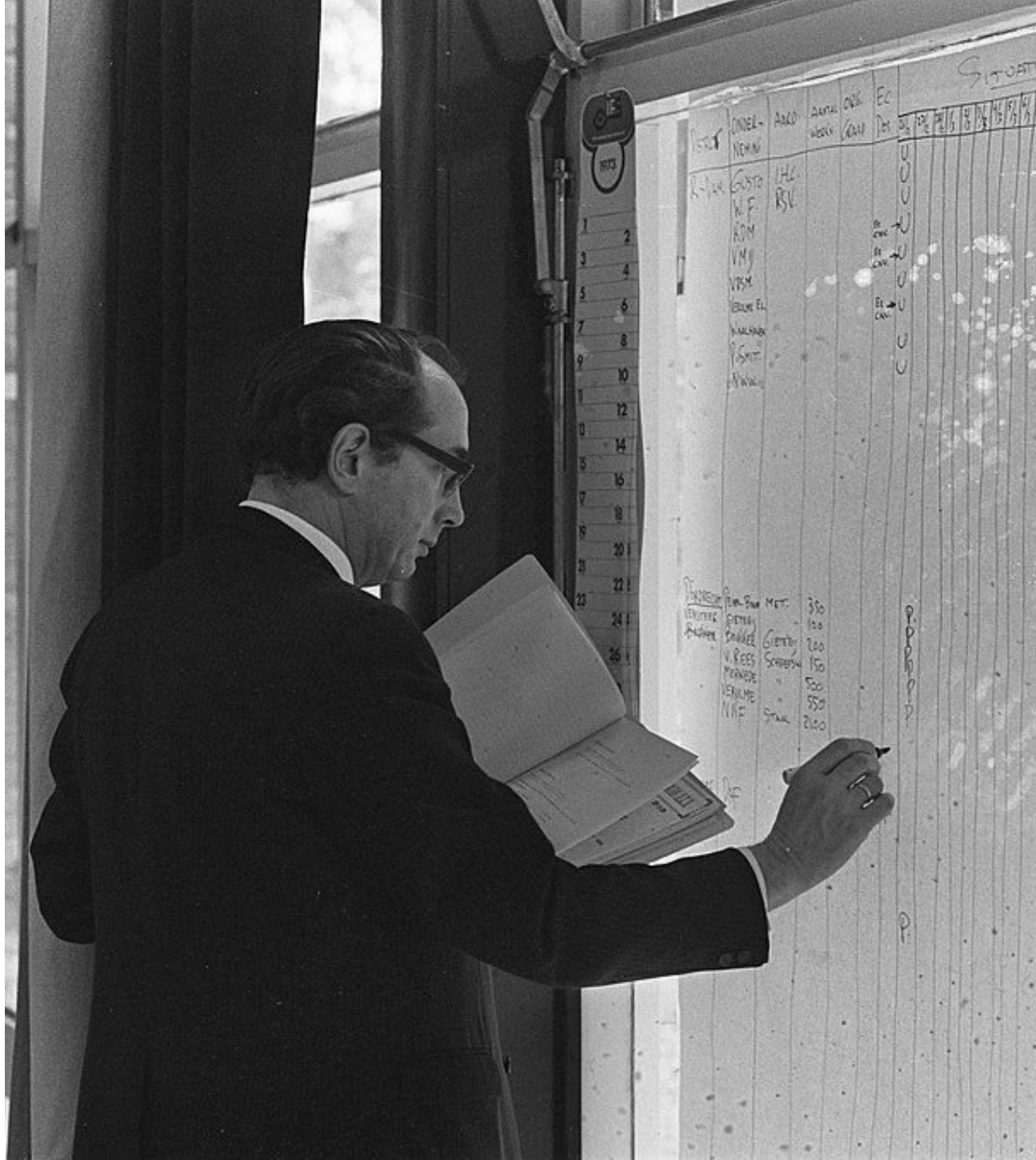


RDF Graph

PREMIS OWL + Schema.org/bibframe

Tides of wisdom

aka lessons learned and the way forward



Lessons learned

A meemoo-SIP specification is not enough for our partners

- non-technical manuals for creating specific meemoo-SIP
- support in metadata mappings

Information about E-ARK is scattered over websites and specs

- avoidable breaking changes & custom tooling
- too little contact with the E-ARK community

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

The way forward

Reducing the complexity of current meemoo-SIP

Bridging the E-ARK gap

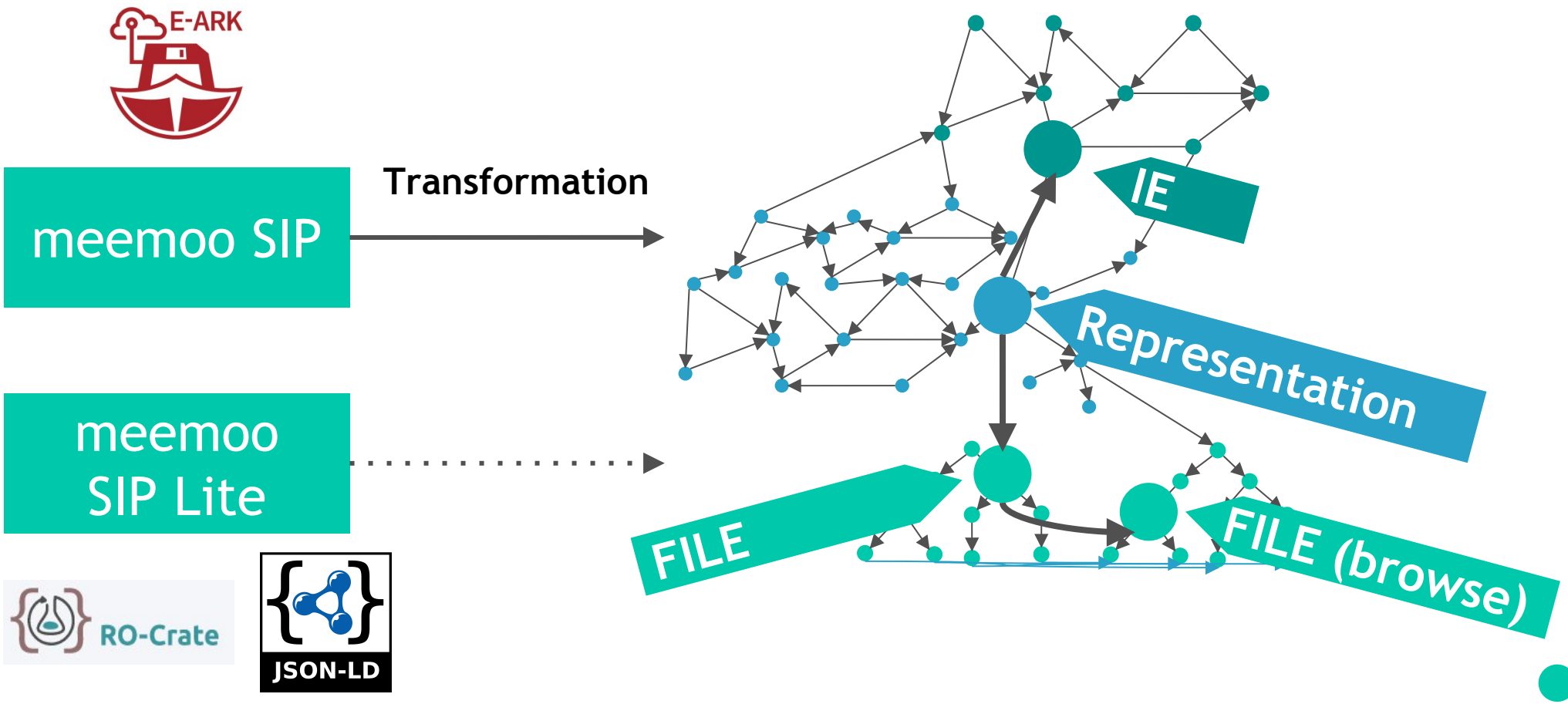
Usage of E-ARK tooling and contributing

Specific documentation for users

Thinking about a SIP lite using JSON-LD

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

The way forward (2)



meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Thanks for listening!

Miel Vander Sande (miel.vandersande@meemoo.be)
Mattias Poppe (mattias.poppe@meemoo.be)
Matthias Priem (matthias.priem@meemoo.be)
Milan Valadou (milan.valadou@meemoo.be)

meemoo
VLAAMS INSTITUUT VOOR HET ARCHIEF

Vlaanderen
verbeelding werkt