

TRUSTWORTHY PRESERVATION AND ACCESS METADATA

Using Distributed Ledger Technology (DLT)

PV2023 Conference 2–4 May 2023, CERN



This project has received funding from the Connecting Europe Facilities (CEF)
Telecom programme by the agreement INEA/CEF/ICT/A2020/2397190
(Action No: 2020-EU-IA-0185)



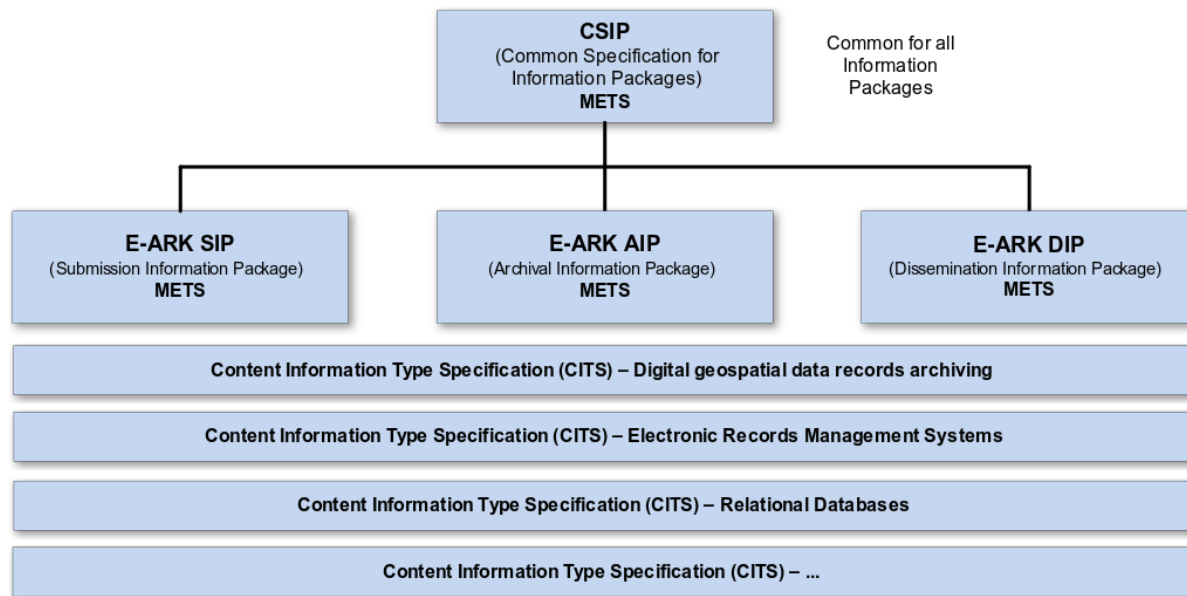
OVERVIEW

- Concept, Principles, and Use Case
- Stakeholder and Collections
- eArchiving Use Case
- Managing and Persisting Resource Access Information
- Blockchain-verifiable Resource Access Metadata



E-ARK STANDARDS & EUROPEAN BLOCKCHAIN SERVICE INFRASTRUCTURE

eArchiving (E-ARK Standards)



European Blockchain Service Infrastructure



Cooperation & Synergies



This project has received funding from the Connecting Europe Facilities (CEF) Telecom programme by the agreement INEA/CEF/ICT/A2020/2397190 (Action No: 2020-EU-IA-0185)



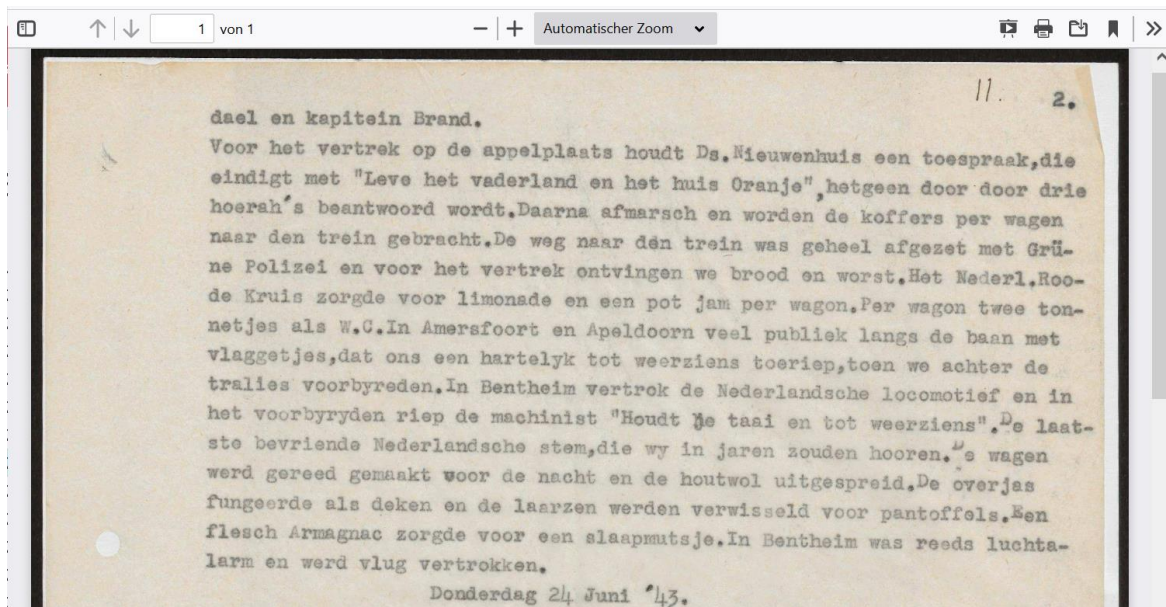
NIOD'S COLLECTIONS

- The **NIOD Institute for War, Holocaust and Genocide Studies** provided several collections to test the concept of **E-ARK/EBSI based access** mechanisms.
 - Archive 244 European **Diaries and Egodocuments**
 - Diaries reporting stories of pain and loss, fear and hunger and moments of levity while facing greatest misery.
 - Archive 250D **Concentration Camps and Prisons**
 - Personal accounts of former prisoners of concentration camps.
 - Archive 250K **Concentration Camps outside the Netherlands.**
 - Records by private persons, interest groups of former inmates and organizations with the task to determine the fate of missing persons.
 - Archive 817 Ravensbrück **Women's Concentration Camp**
 - The collection includes personal Statements of former prisoners of Ravensbrück.

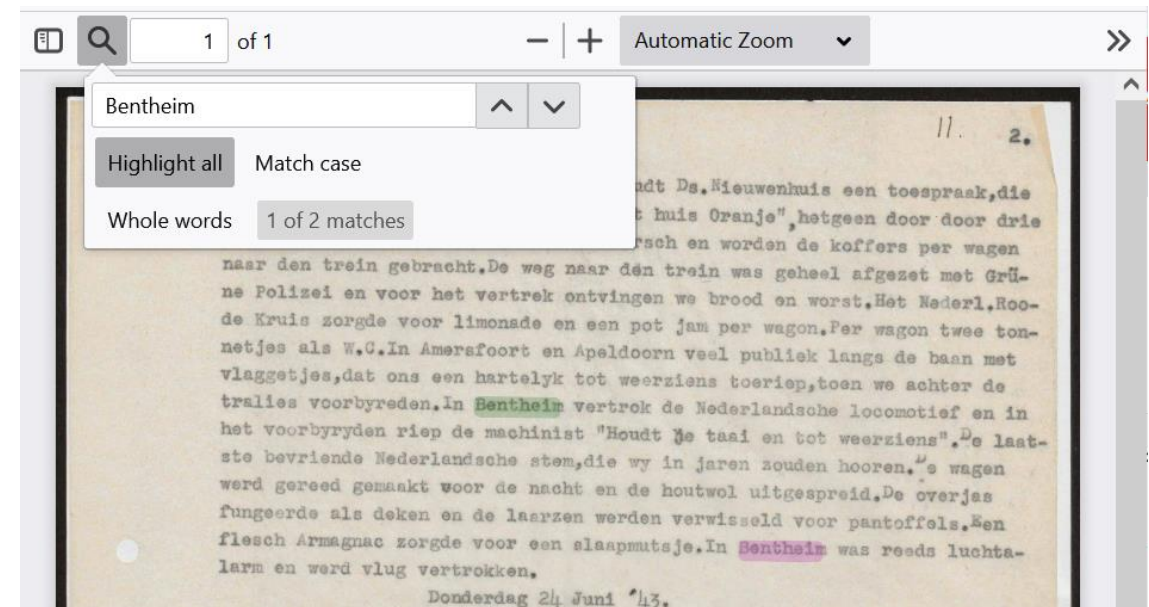


MASTER/DISSEMINATION REPRESENTATION

MASTER Image Representation

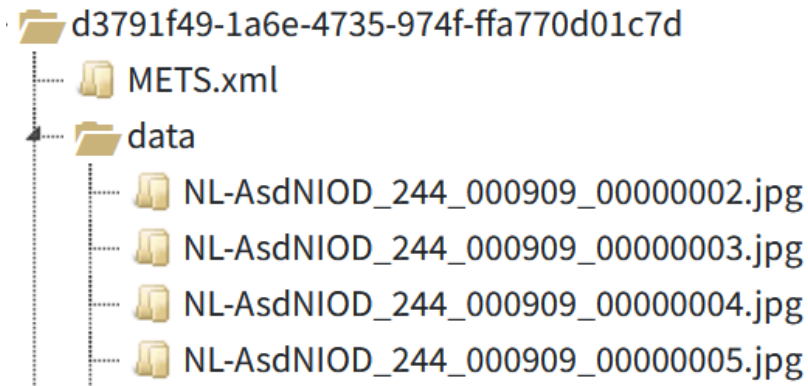


Searchable PDF Representation

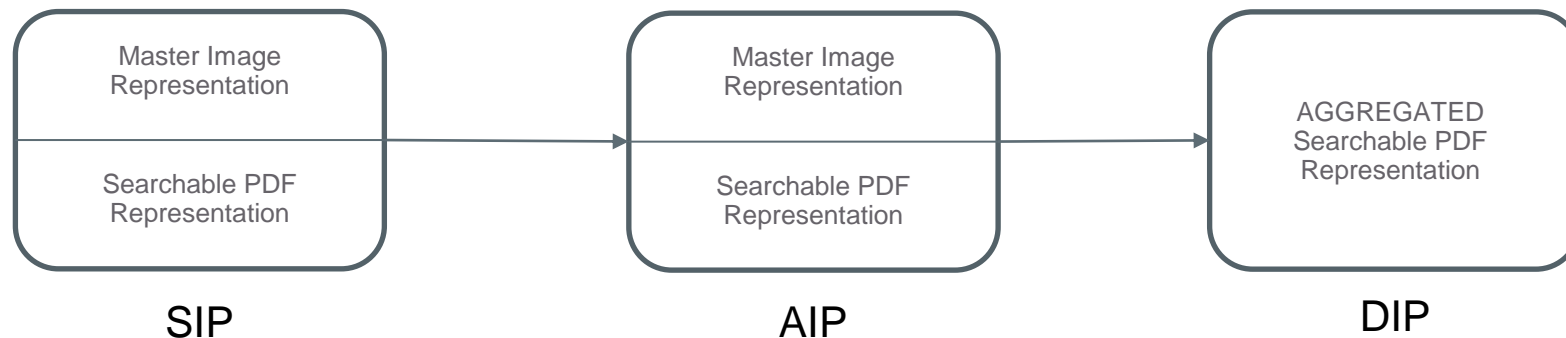
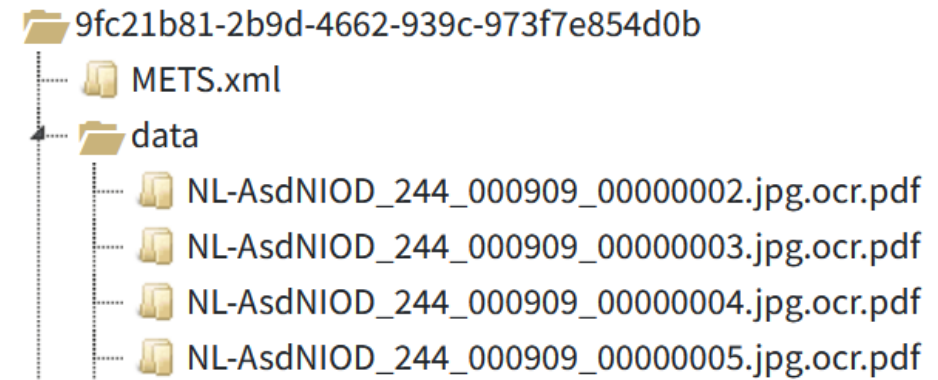


MASTER/DISSEMINATION REPRESENTATION

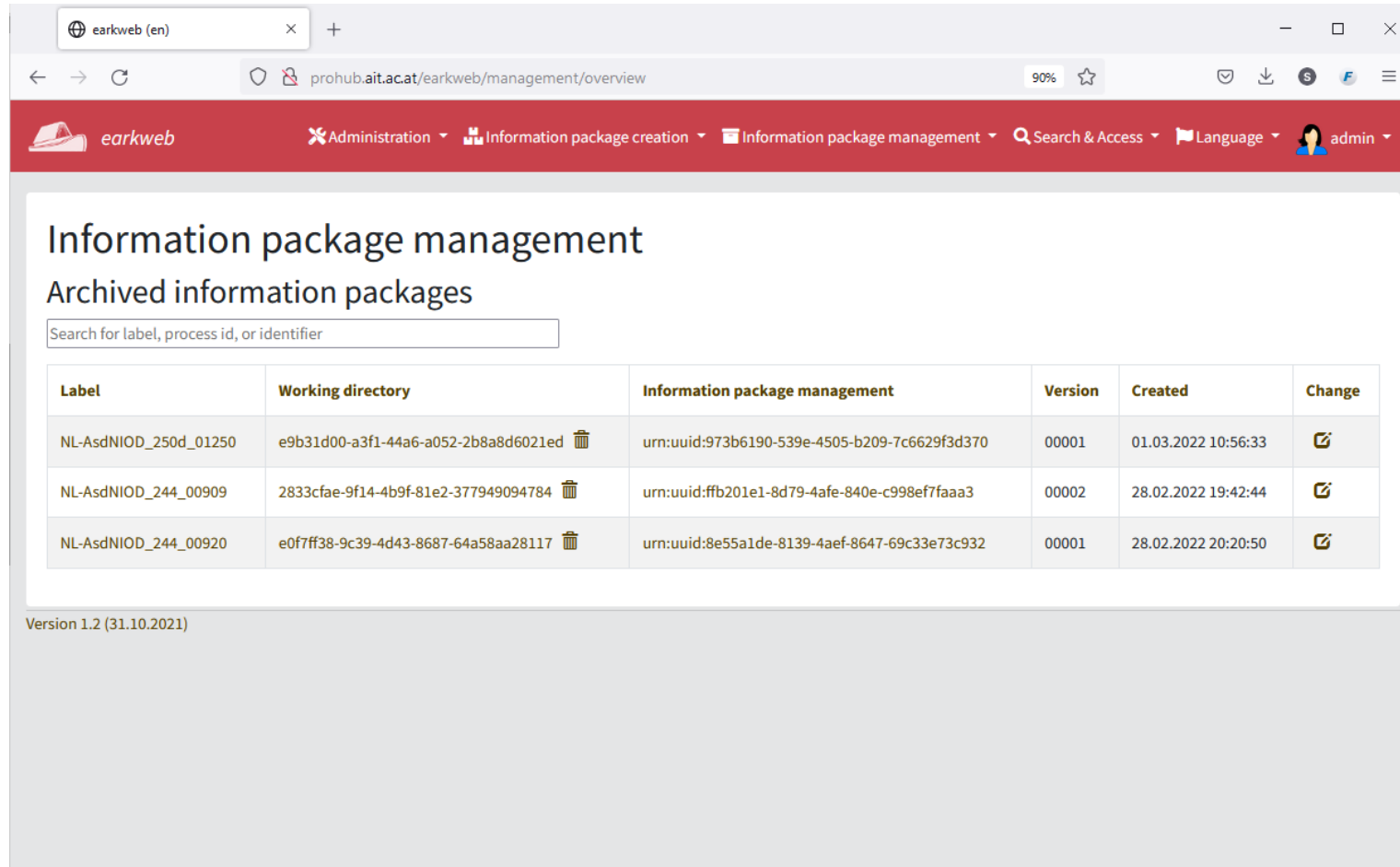
Master Image Representation









Searchable PDF Representation



E-ARK REPOSITORY REFERENCE IMPLEMENTATION

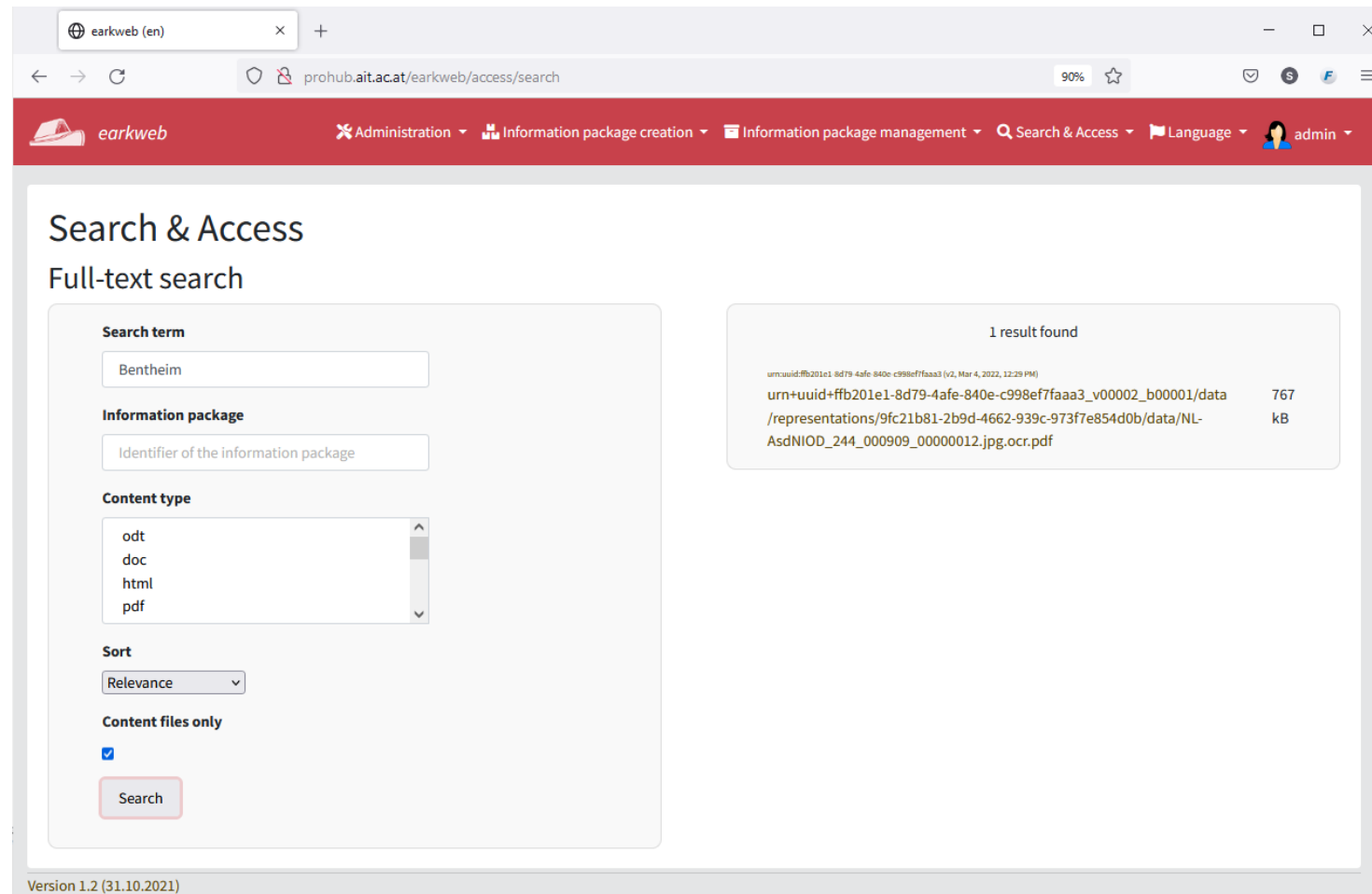


The screenshot shows a web browser window with the URL `prohub.ait.ac.at/earkweb/management/overview`. The page title is "Information package management" and the subtitle is "Archived information packages". There is a search bar with the placeholder text "Search for label, process id, or identifier". Below the search bar is a table with the following data:

Label	Working directory	Information package management	Version	Created	Change
NL-AsdNIOD_250d_01250	e9b31d00-a3f1-44a6-a052-2b8a8d6021ed 	urn:uuid:973b6190-539e-4505-b209-7c6629f3d370	00001	01.03.2022 10:56:33	
NL-AsdNIOD_244_00909	2833cfae-9f14-4b9f-81e2-377949094784 	urn:uuid:ffb201e1-8d79-4afe-840e-c998ef7faaa3	00002	28.02.2022 19:42:44	
NL-AsdNIOD_244_00920	e0f7ff38-9c39-4d43-8687-64a58aa28117 	urn:uuid:8e55a1de-8139-4aef-8647-69c33e73c932	00001	28.02.2022 20:20:50	

Version 1.2 (31.10.2021)

SEARCH



The screenshot shows a web browser window with the URL `prohub.ait.ac.at/earkweb/access/search`. The page title is "Search & Access" and the sub-section is "Full-text search". The search term entered is "Bentheim". The search results show one result found, with the following details:

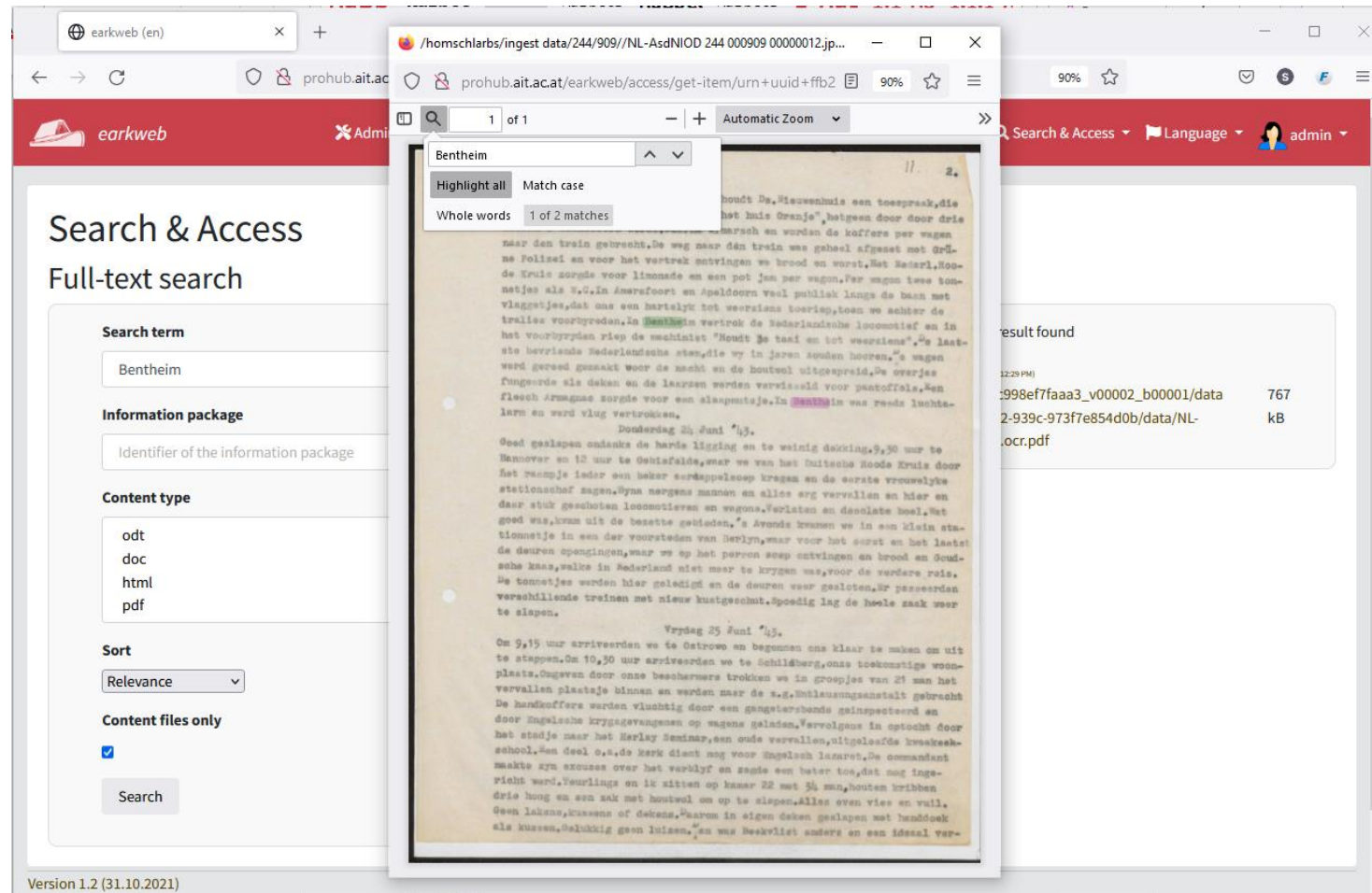
Search term	Information package	Content type	Sort	Content files only
Bentheim	Identifier of the information package	odt doc html pdf	Relevance	<input checked="" type="checkbox"/>

The search results table shows the following details for the single result found:

1 result found
<code>urn:uuid:ffb201e1-8d79-4afe-840e-c998ef7faaa3 (v2, Mar 4, 2022, 12:29 PM)</code>
<code>urn+uuid+ffb201e1-8d79-4afe-840e-c998ef7faaa3_v00002_b00001/data</code> 767
<code>/representations/9fc21b81-2b9d-4662-939c-973f7e854d0b/data/NL-</code> kB
<code>AsdNIOD_244_000909_00000012.jpg.ocr.pdf</code>

Version 1.2 (31.10.2021)

SEARCH



The screenshot displays the earkweb search interface. On the left, the 'Search & Access' sidebar is visible, containing a search form with the following fields and options:

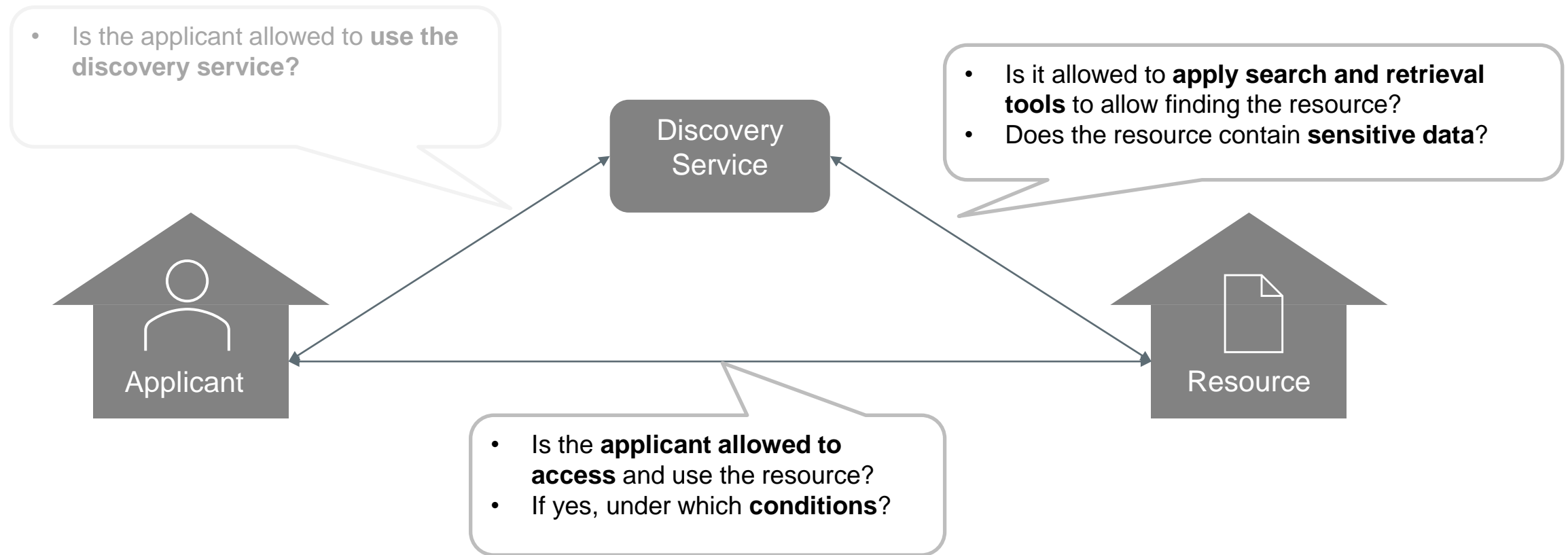
- Search term:** Bentheim
- Information package:** Identifier of the information package
- Content type:** odt, doc, html, pdf
- Sort:** Relevance
- Content files only:**

The main search results area shows a single result for the document 'Bentheim'. A search bar at the top of the document viewer indicates '1 of 1' matches. A search bar within the document viewer shows 'Bentheim' with '1 of 2 matches' highlighted. The document text is in Dutch and discusses historical events, including the arrival of a train and the search for a person named Bentheim.

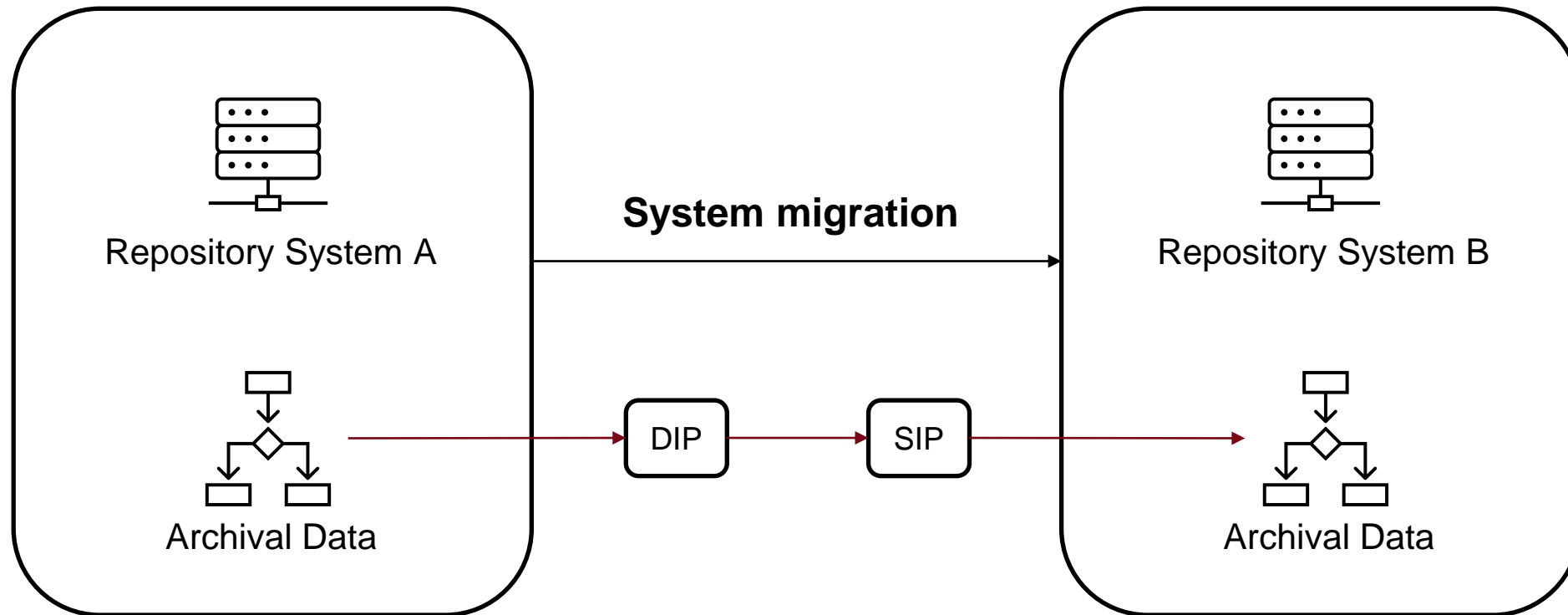
At the bottom of the interface, the version number 'Version 1.2 (31.10.2021)' is displayed.



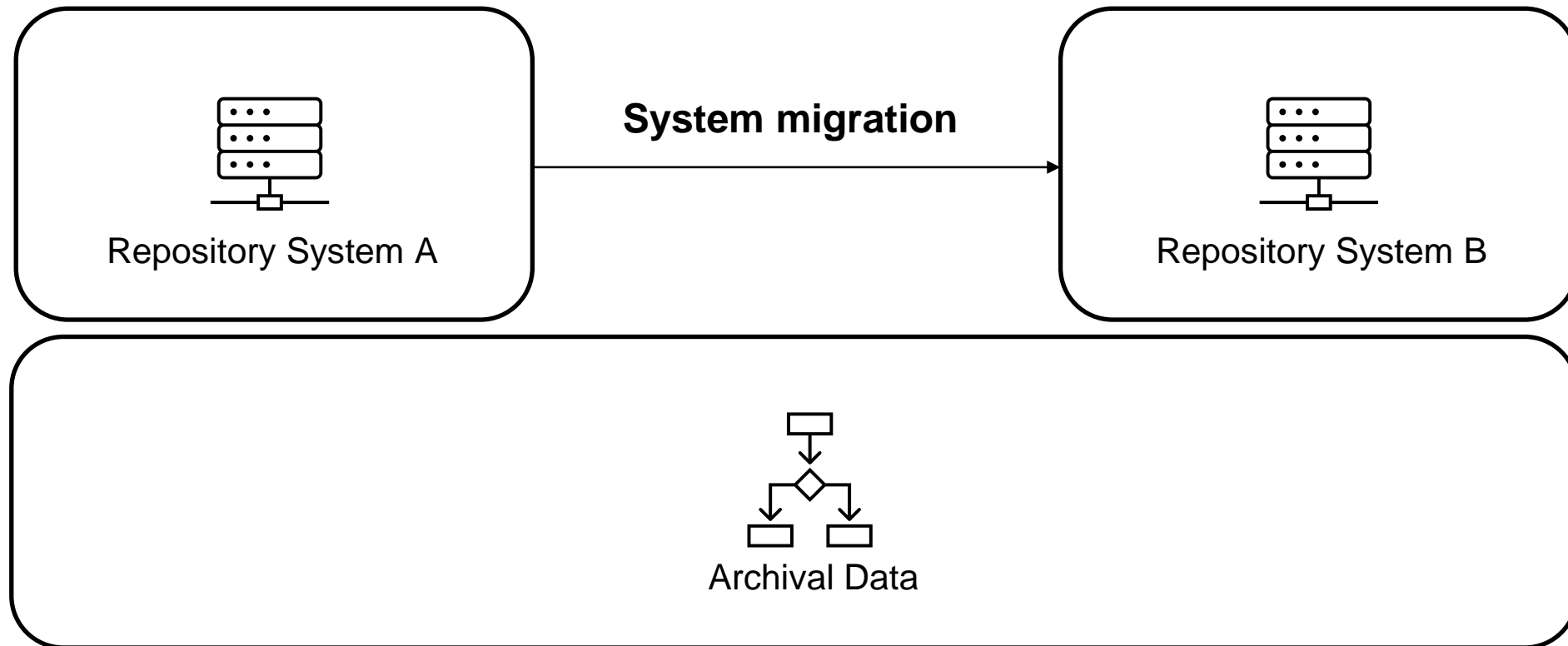
A BASIC AUTHORISATION MODEL



SYSTEM MIGRATION - STILL OFTEN LIKE THIS



AND NOT LIKE THIS?



The amount of data is growing at a fast pace! Moving data will become more and more difficult!

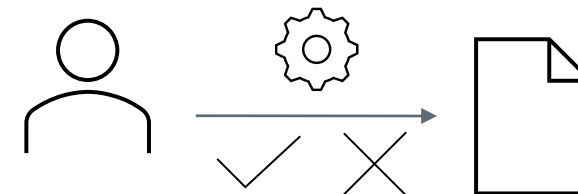
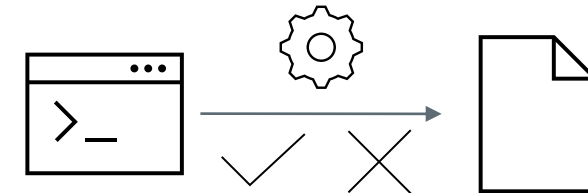
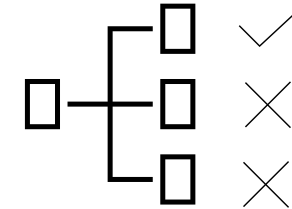
MOTIVATION

Why do we want to include authorisation metadata in the information package?

- *Problem:* Re-configuring authorization settings in a new system is labour-intensive and time-consuming.
- *Approach:* Authorization metadata and licenses remain valid after system migration – signatures, and licenses are interpreted by the new system

REQUIREMENTS

1. It must be possible to define *rights statements* on **information package, representation, or file level**.
2. It must be possible to define *rights statements* related to the application of **software agents** enabling information retrieval, such as Optical Character Recognition (OCR) or full-text indexing software, for example.
3. It must be possible to define *rights statements* regarding information **applicants**.



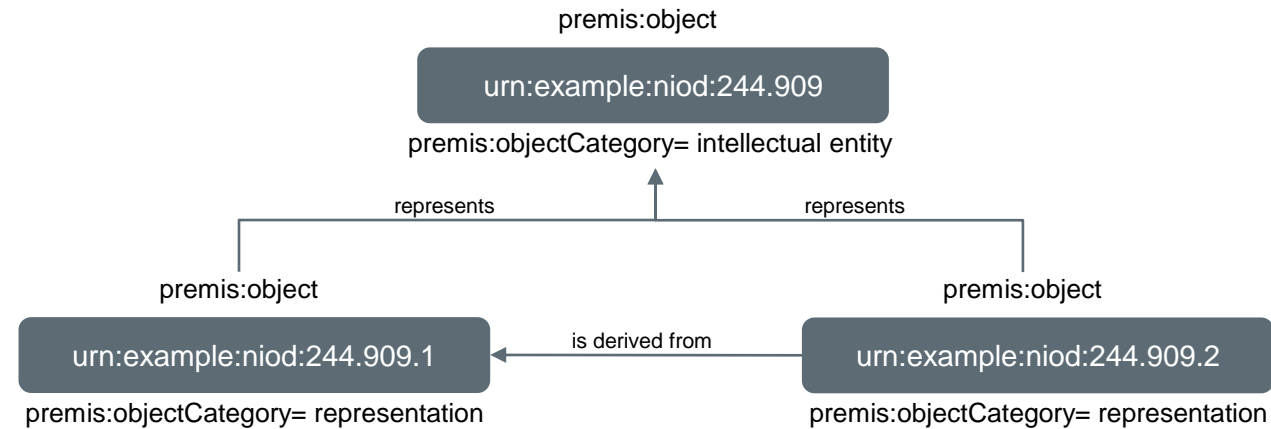
REQUIREMENT 1: OBJECT LEVELS

premis:object

urn:example:niod:244.909

premis:objectCategory= intellectual entity

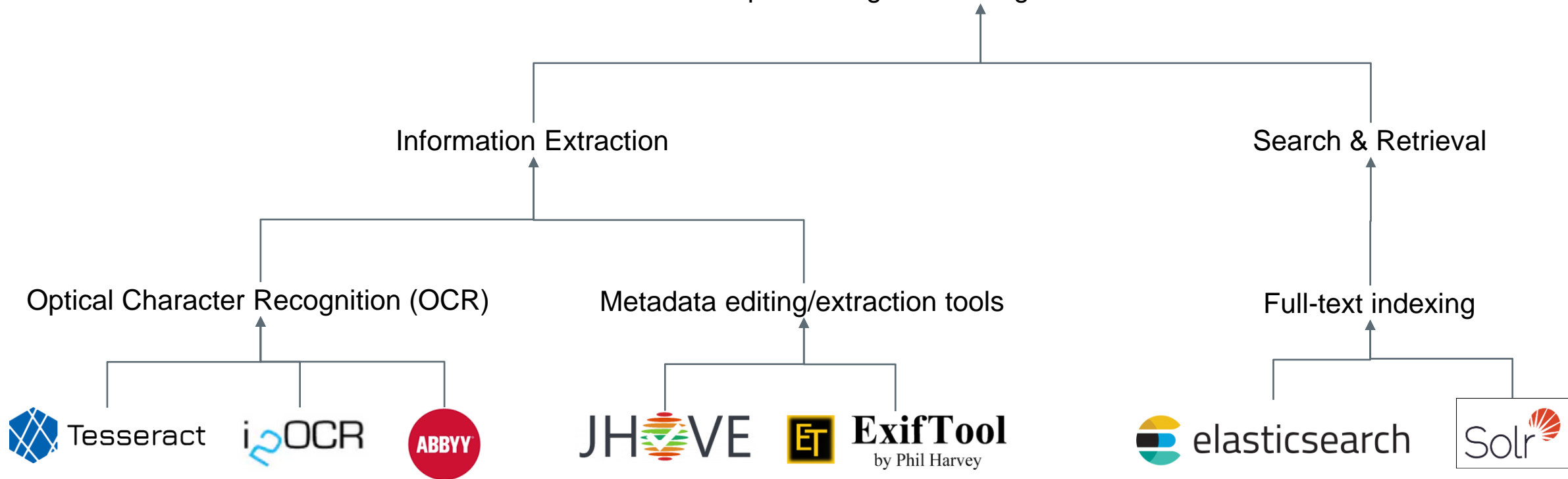
REQUIREMENT 1: OBJECT LEVELS



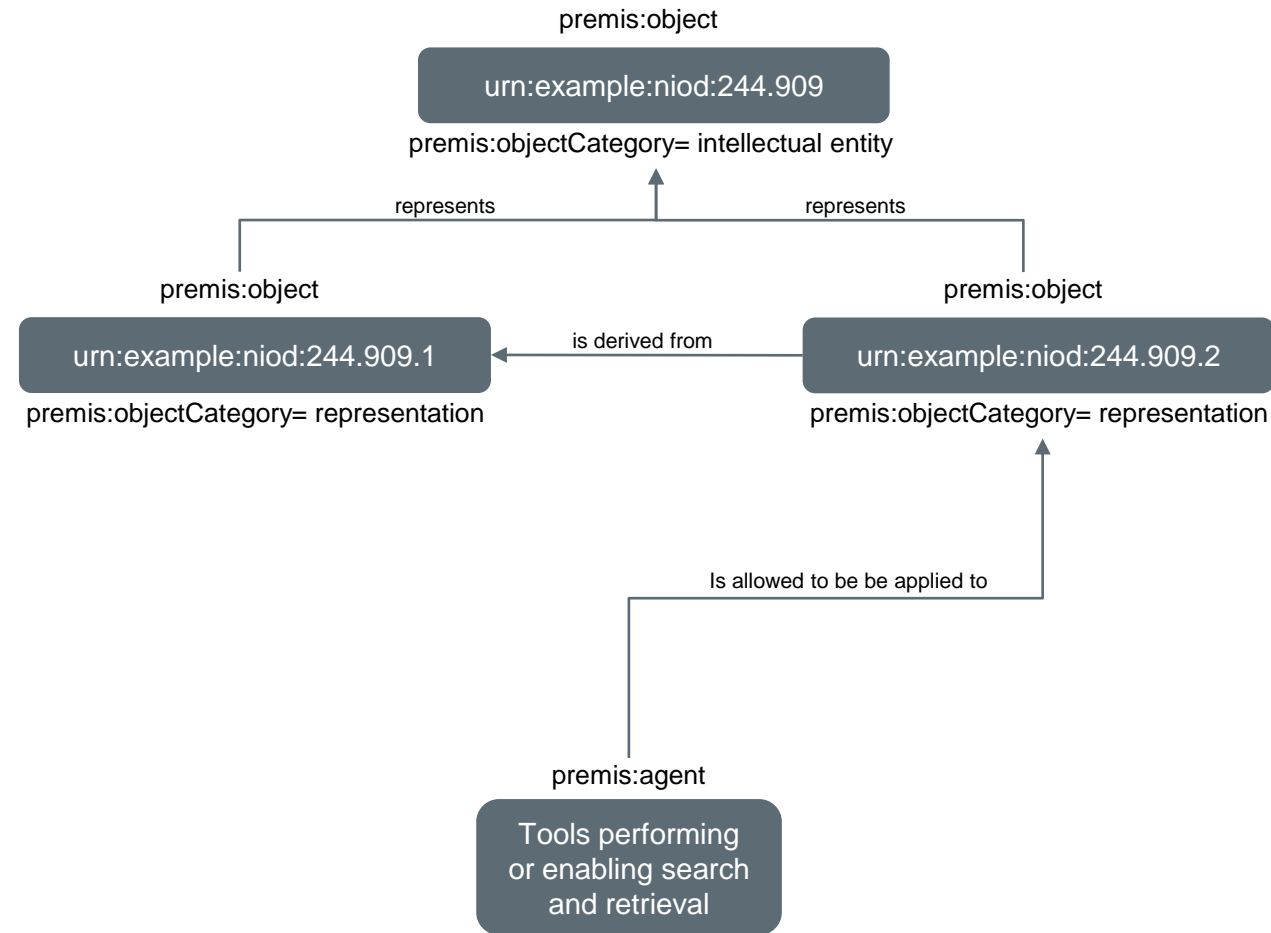
PREMIS Objects for the information package, representation, or file level can be defined and therefore rights statements can relate to the objects at these levels.

REQUIREMENT 2: SOFTWARE AGENTS

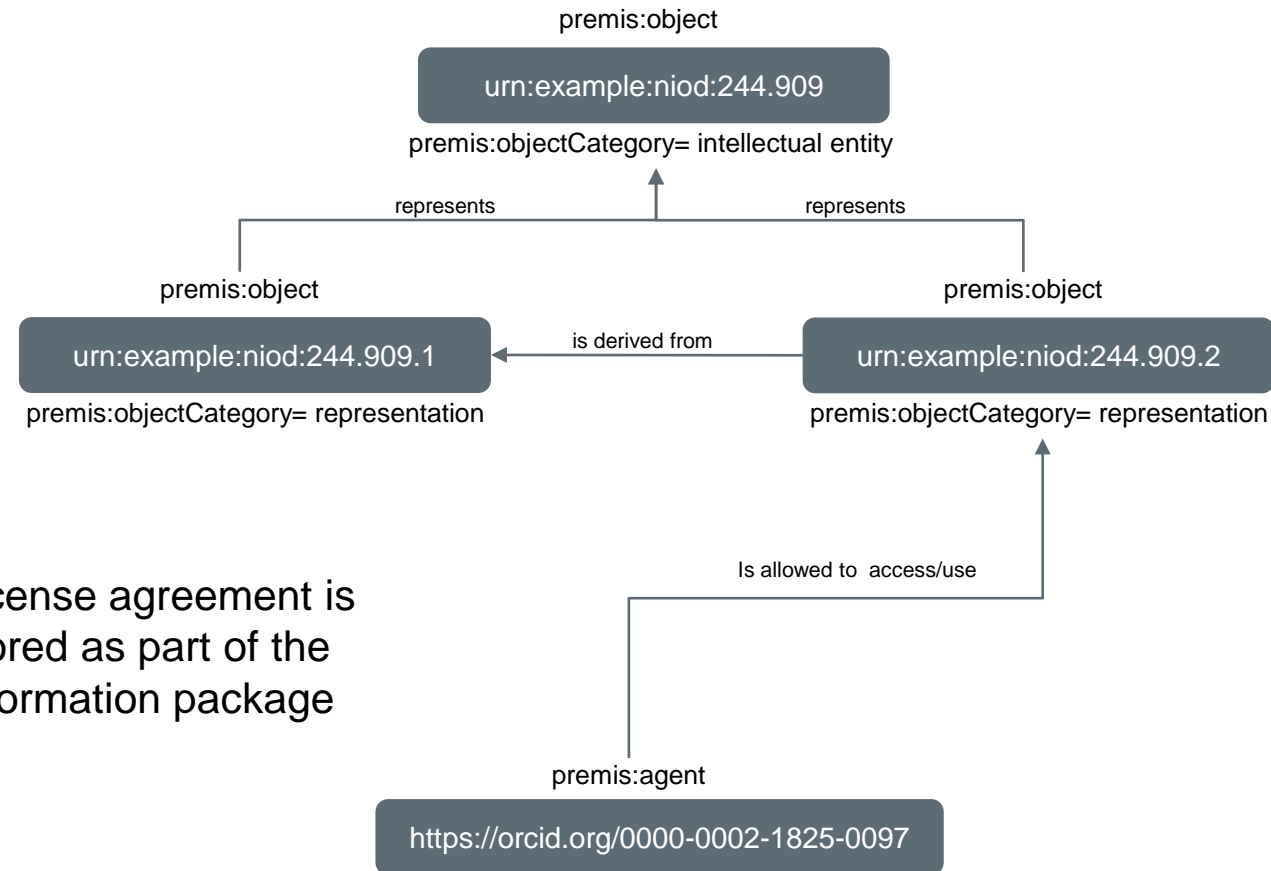
Tools performing or enabling search and retrieval

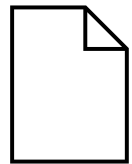


REQUIREMENT 2: SOFTWARE AGENTS

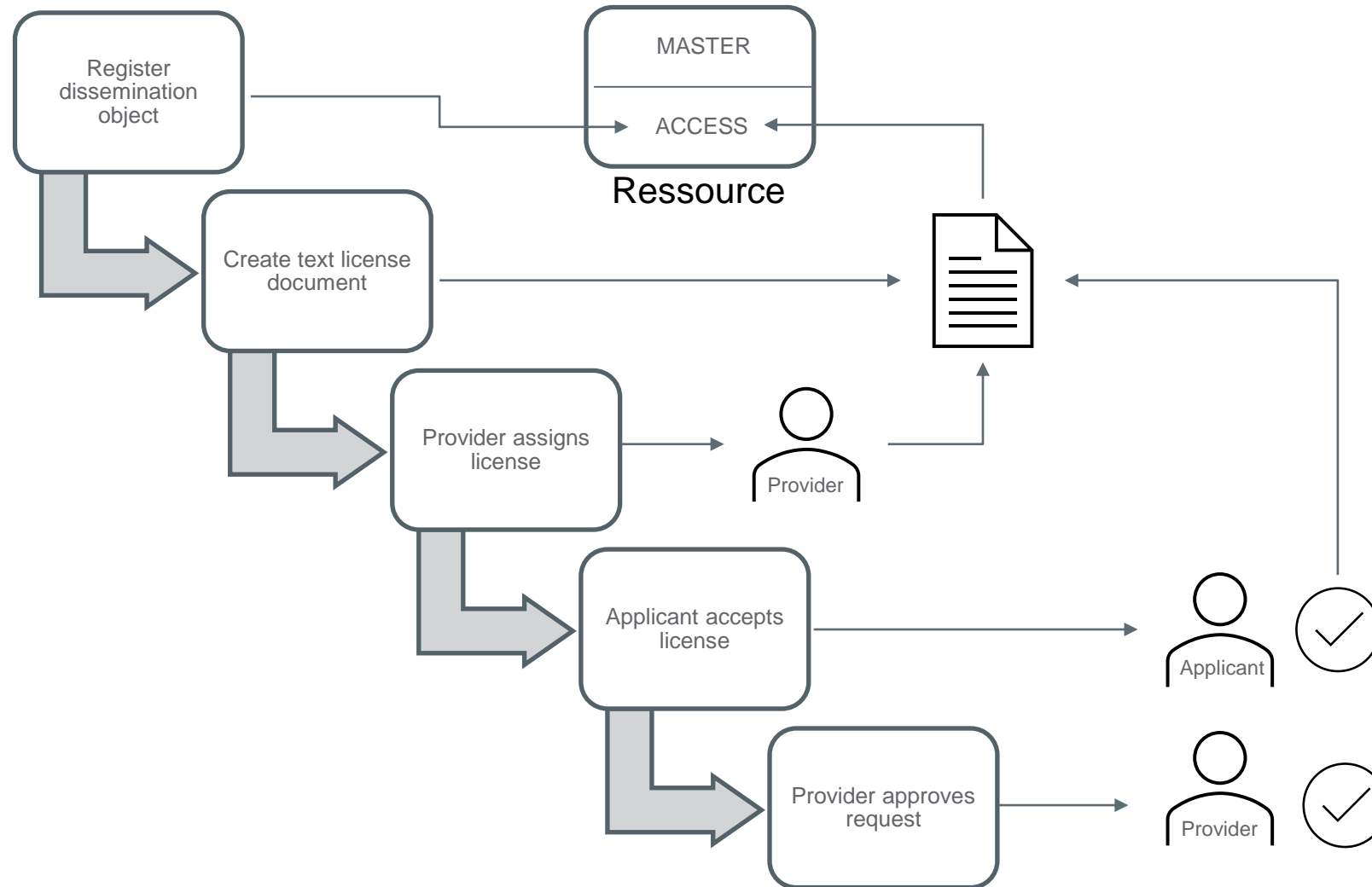


REQUIREMENT 3: APPLICANT



 License agreement is stored as part of the information package

BLOCKCHAIN OBJECT REGISTRATION AND LICENSE AGREEMENT PROCESS



EUROPEAN BLOCKCHAIN SERVICES INFRASTRUCTURE (EBSI)

What is EBSI?

The European Blockchain Services Infrastructure (EBSI) is a market-friendly distributed blockchain network based on open standards and transparent governance model.



Source: <https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/What+is+ebsi>



This project has received funding from the Connecting Europe Facilities (CEF) Telecom programme by the agreement INEA/CEF/ICT/A2020/2397190 (Action No: 2020-EU-IA-0185)



EBSI USE CASES

Identity

Diploma

Social security


Document traceability

(Coming soon)



Document traceability

The **Document traceability Use Case** aims to leverage blockchain to create trusted digital audit trails, automate compliance checks – for example in time-sensitive processes – and to prove data integrity. The intention is also to serve as generic registration and traceability capabilities for other EBSI Use cases.

Documentation 

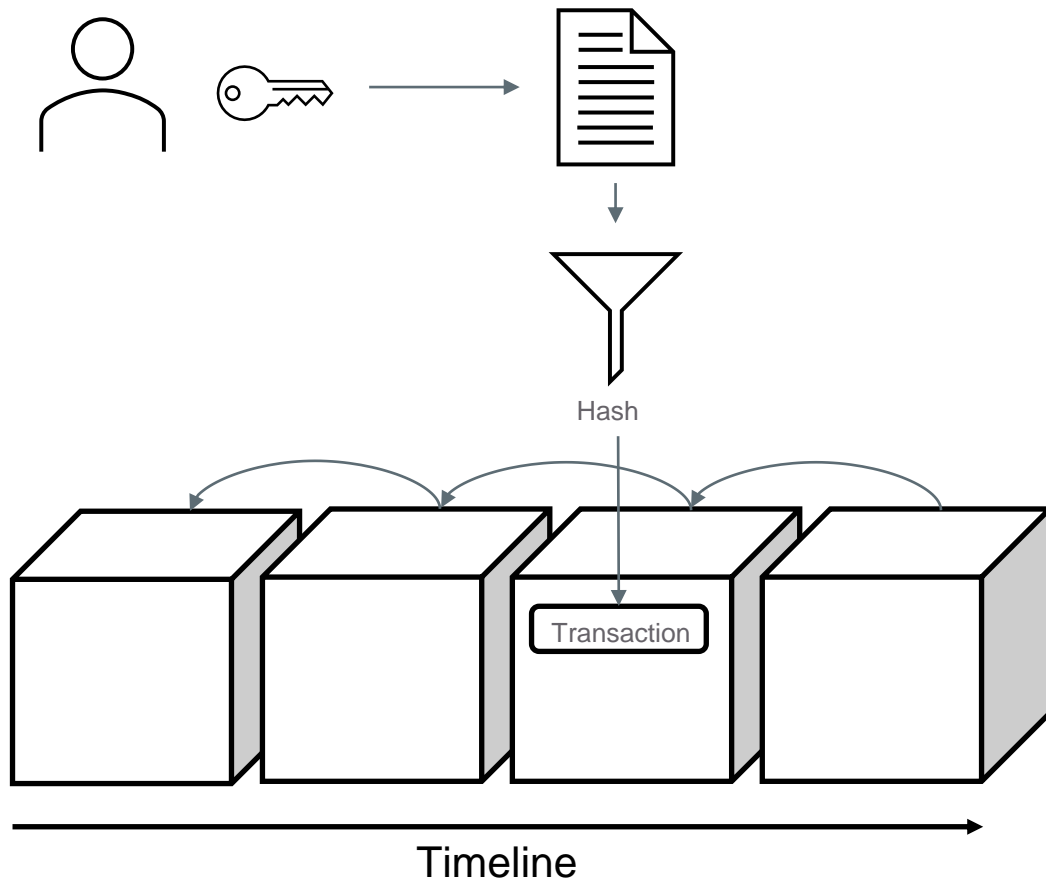
Source: <https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Use+cases>



This project has received funding from the Connecting Europe Facilities (CEF)
Telecom programme by the agreement INEA/CEF/ICT/A2020/2397190
(Action No: 2020-EU-IA-0185)

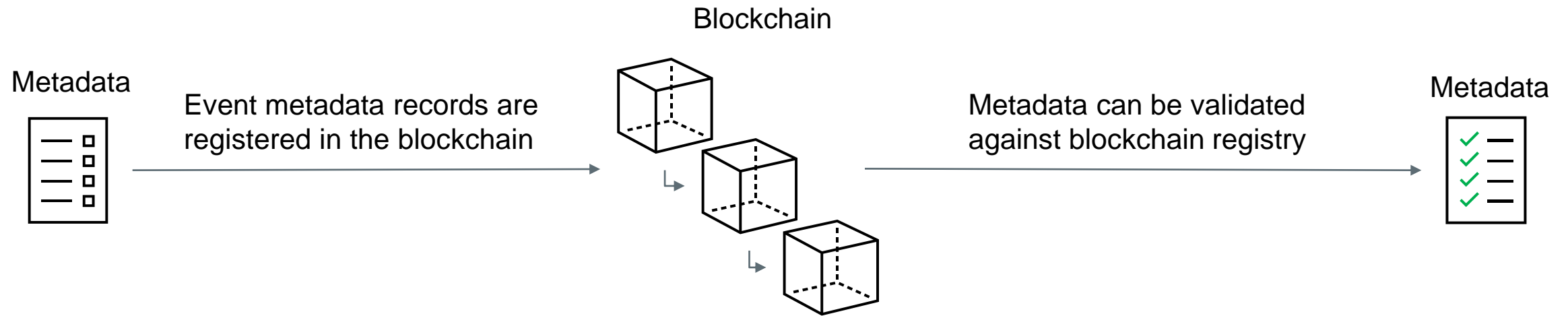


BLOCKCHAIN NOTARY



- Fingerprint (hash) of document(s) and optionally signatures is stored
 - impossible to derive the original document from the hash value only
- Transaction stores the hash value with a timestamp when it was performed
- The hash value may represent a single document or multiple documents
- Documents may be signed or hashed together with signature files
- Proof of existence and (optionally) proof of authenticity

BLOCKCHAIN BACKED METADATA



- Integrity and authenticity
 - Protection against manipulation
- } of resource access and authorization metadata

PROTOTYPE IMPLEMENTATION

- Prototype published at:
 - <https://github.com/E-ARK-Software/blockchain-notary-poc>



THANK YOU!

SVEN SCHLARB

Scientist

Data Science & Artificial Intelligence
Center for Digital Safety & Security

AIT Austrian Institute of Technology GmbH

Giefinggasse 4 | 1210 Vienna | Austria

sven.schlarb@ait.ac.at



This project has received funding from the Connecting Europe Facilities (CEF)
Telecom programme by the agreement INEA/CEF/ICT/A2020/2397190
(Action No: 2020-EU-IA-0185)

