



MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE

Our Institute



Rare Book Collection, Max Planck Institute for the History of Science



What kind of Research Data are we dealing with in the Humanities?



What kind of Research Data are we dealing with in the Humanities?

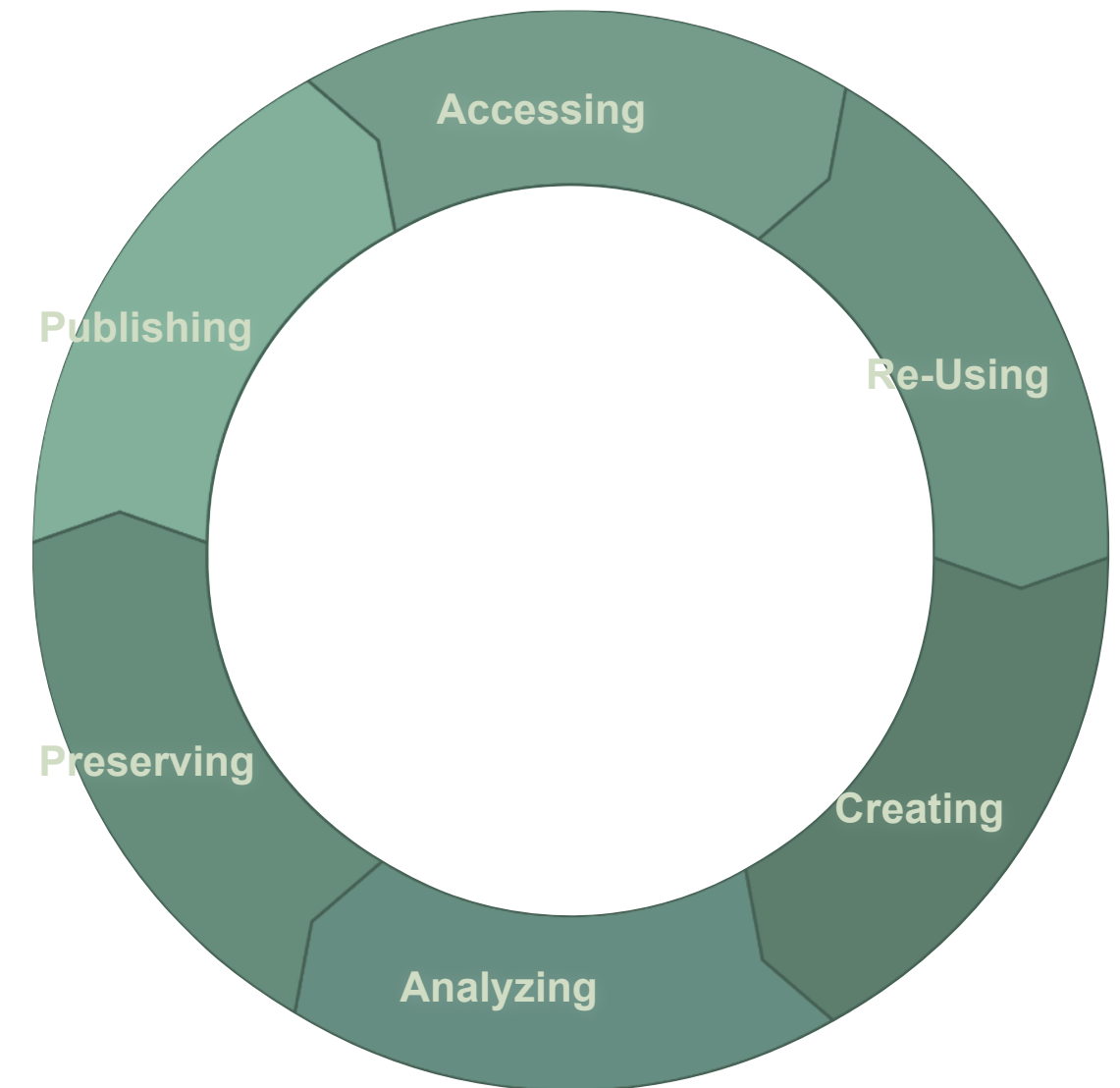




Challenges

Doing Research in the Digital Age

- **Research data: new formats of research output**
- **Lack of use of standards**
- **Incompatible data**
- **Data tied to old software**
- **Vanishing data**
- **Disruption of the research lifecycle**



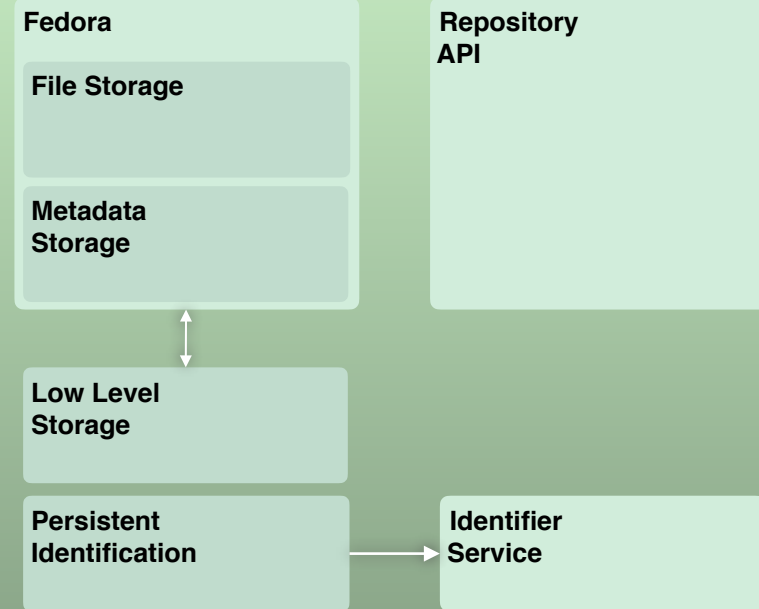


Our Philosophy

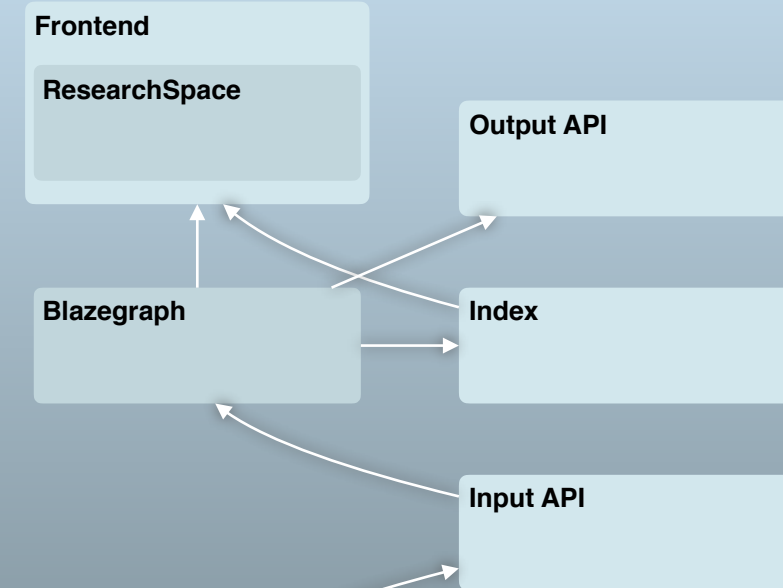
- **Separate data from software**
- **Design for end-of-life**
- **Use dedicated systems and add harmonised access**

Architecture

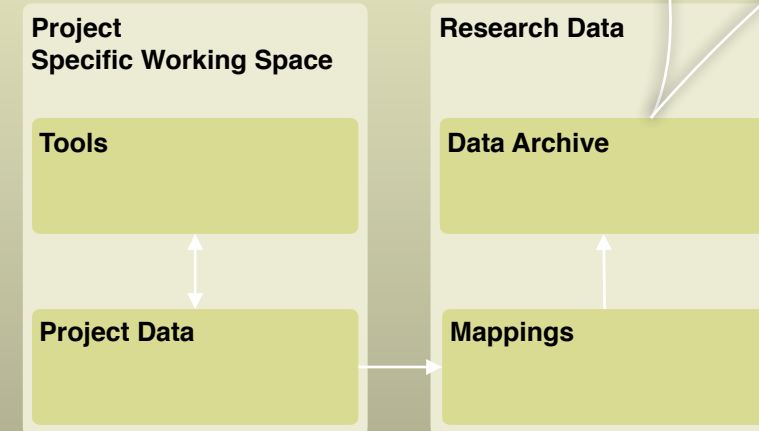
Repository Module



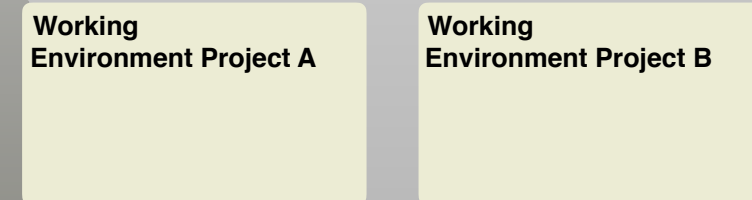
Knowledge Graph



Working Environment



Working Environments



Backup



MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE



Tech Stack

REPOSITORY

- **Fedora 5**
- **CDSTAR**
- **Solr**
- **Cantaloupe Image Server**
- Backend API (Java)
- Admin Frontend (React)

INGEST

- **Goobi**

DATA ARCHIVE

- **DARIAH Repository**

KNOWLEDGE GRAPH

- **ResearchSpace**
- **Blazegraph**

MAPPING

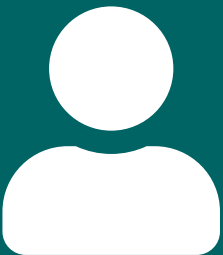
- **X3ML Mapping Engine**
- **3M Mapping Tool**

“GLUE”

- **GitLab (code base and deployment)**
- **RabbitMQ (message broker)**
- **KeyCloak (authorisation)**



Workflow Digitisation



Goobi - 30

grobi.mpg-berlin.mpg.de/grobi/vtask_all.xhtml

Dashboard My tasks Workflow Administration Management

Current user: 1 10:55 AM Wednesday, 06. November 2019

Processes

Dashboard > Processes

Processes

Items: 39

ID	Process title	Status	Project	Action
139	> 1025141482		Beispielprojekt	/ # Q +
140	> 1030202701		Beispielprojekt	/ # Q +
141	> 1039202051		Beispielprojekt	/ # Q +
138	> 1041240475		Beispielprojekt	/ # Q +
137	> 1047189569		Beispielprojekt	/ # Q +
149	> 1299206037_1808		Beispielprojekt	/ # Q +
150	> 1299206037_1809		Beispielprojekt	/ # Q +
151	> 1299206037_1810		Beispielprojekt	/ # Q +
152	> 1299206037_1812		Beispielprojekt	/ # Q +
153	> 1299206037_1813		Beispielprojekt	/ # Q +

Previous page Page 1 of 4 Next page

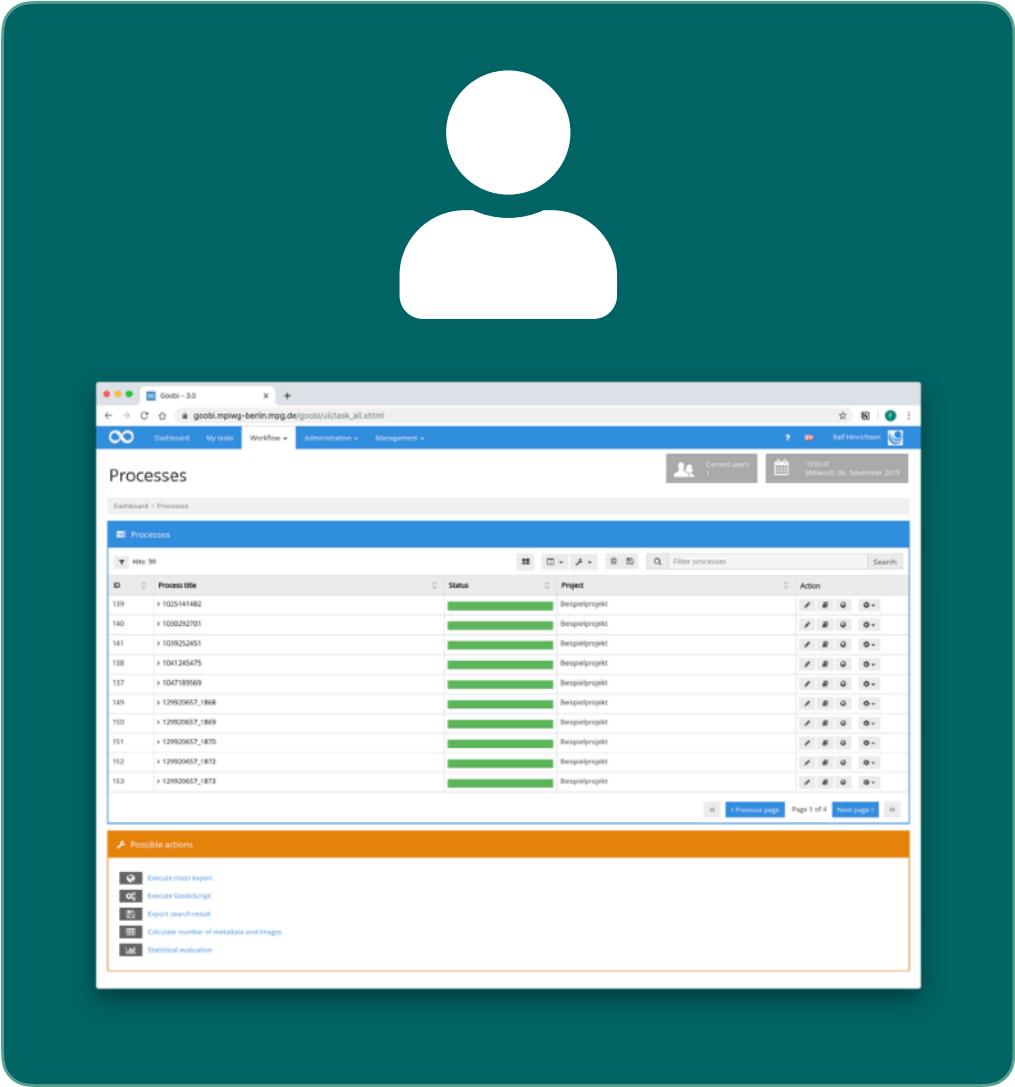
Possible actions

- [Execute mass export](#)
- [Execute GoobiImage](#)
- [Export search result](#)
- [Calculate number of metadata and images](#)
- [Statistical evaluation](#)

Goobi

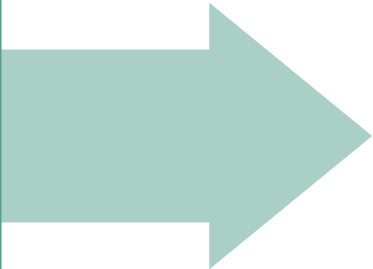


Workflow Digitisation



Goobi

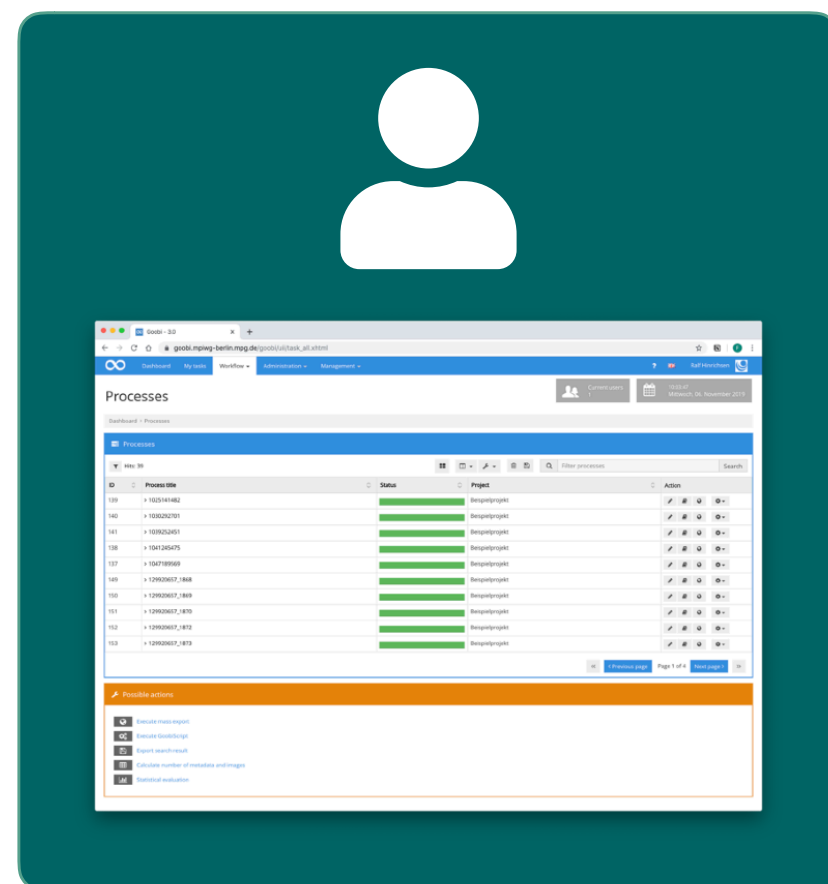
Mets/Mods XML



Media Files



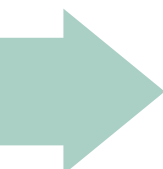
MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE



Goobi



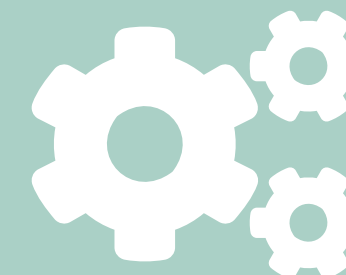
Repository



Repository



Mapping



X3ML Mapping Engine

Repository





MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE

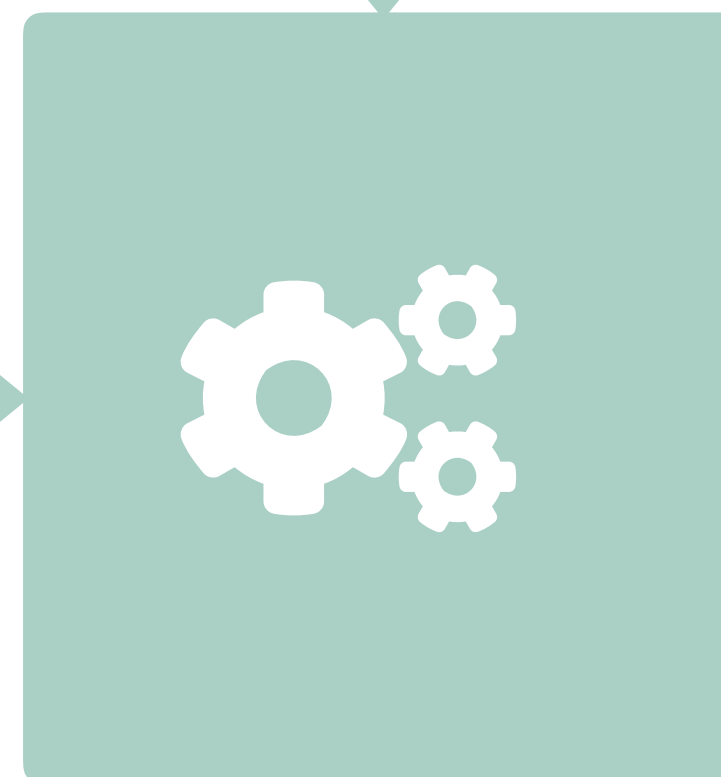
Mapping



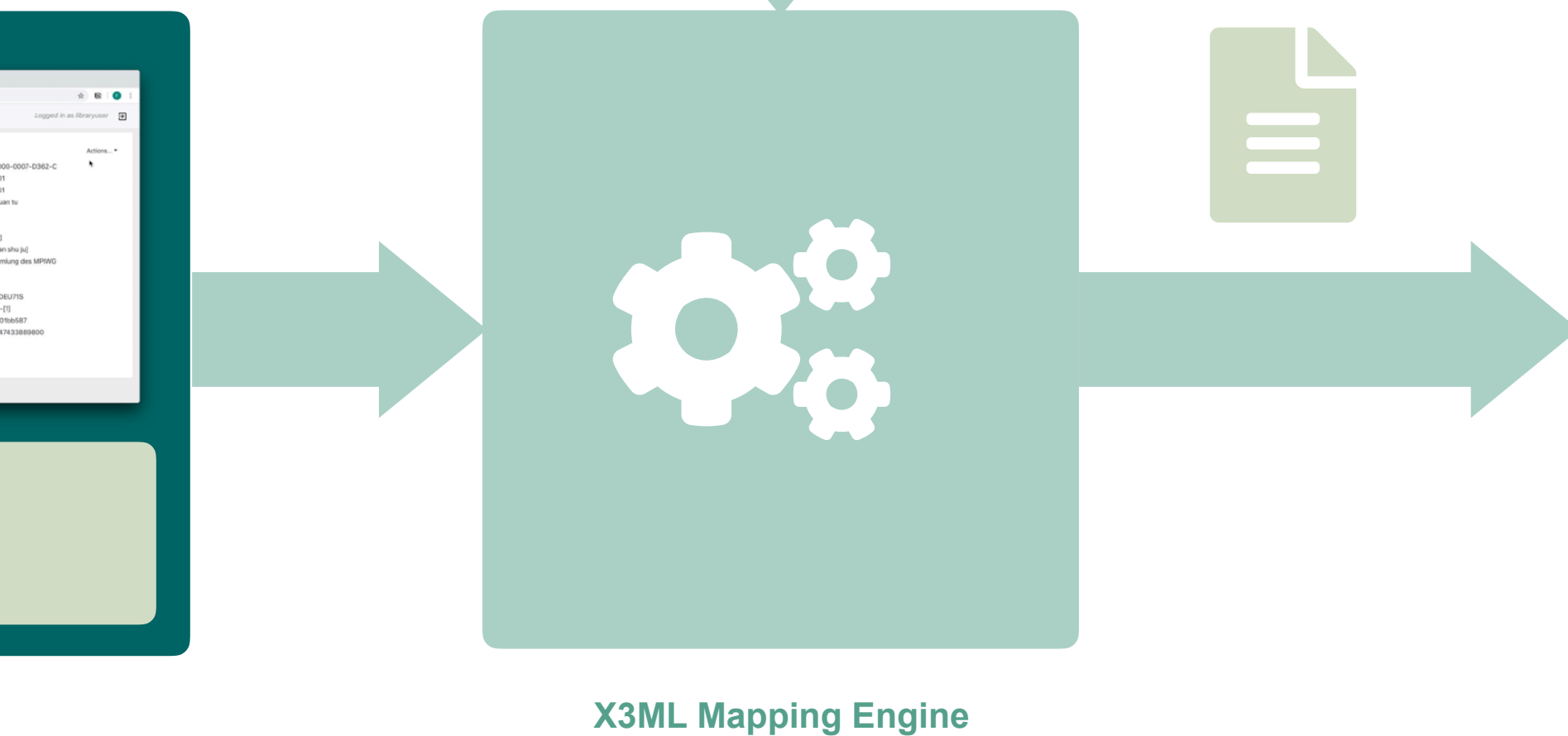
CIDOC/RDF



Repository

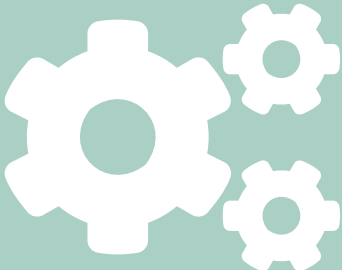


X3ML Mapping Engine





Mapping



X3ML Mapping Engine



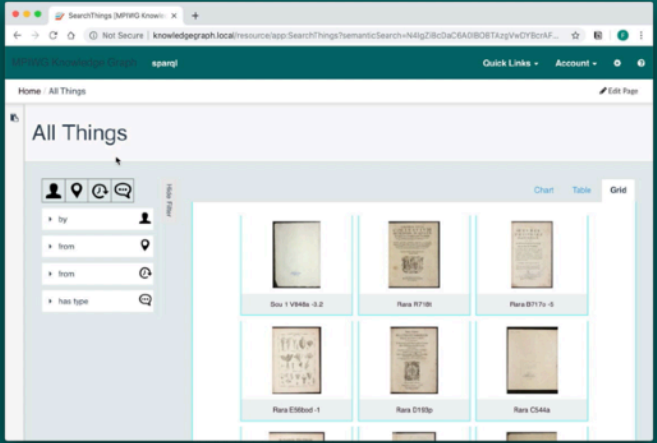
ResearchSpace



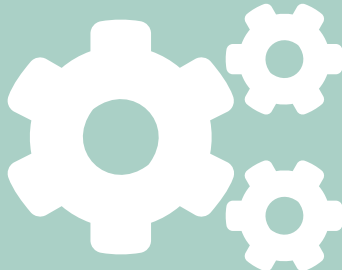
metaphacts



Knowledge Graph

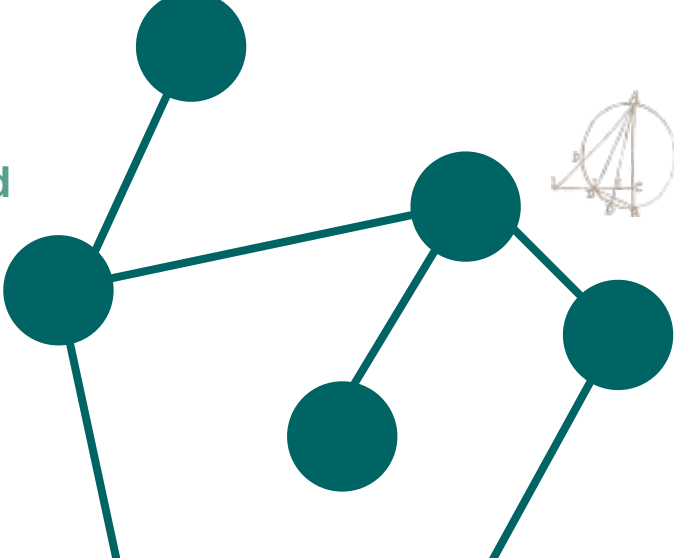


Mapping



X3ML Mapping Engine

LOD Cloud

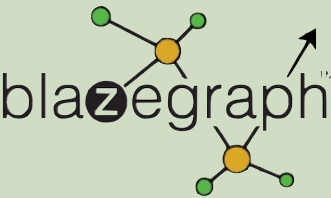


MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE

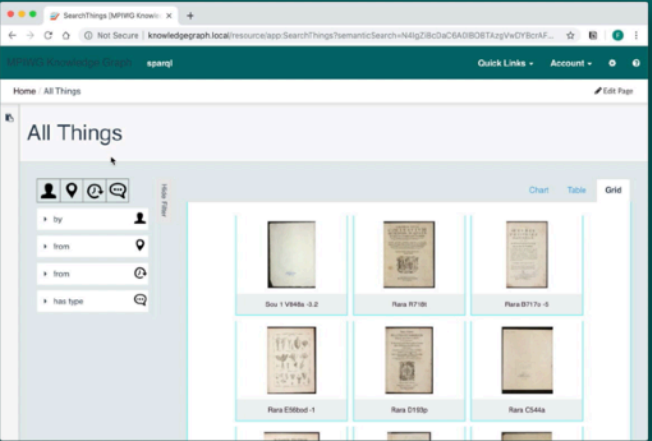
ResearchSpace

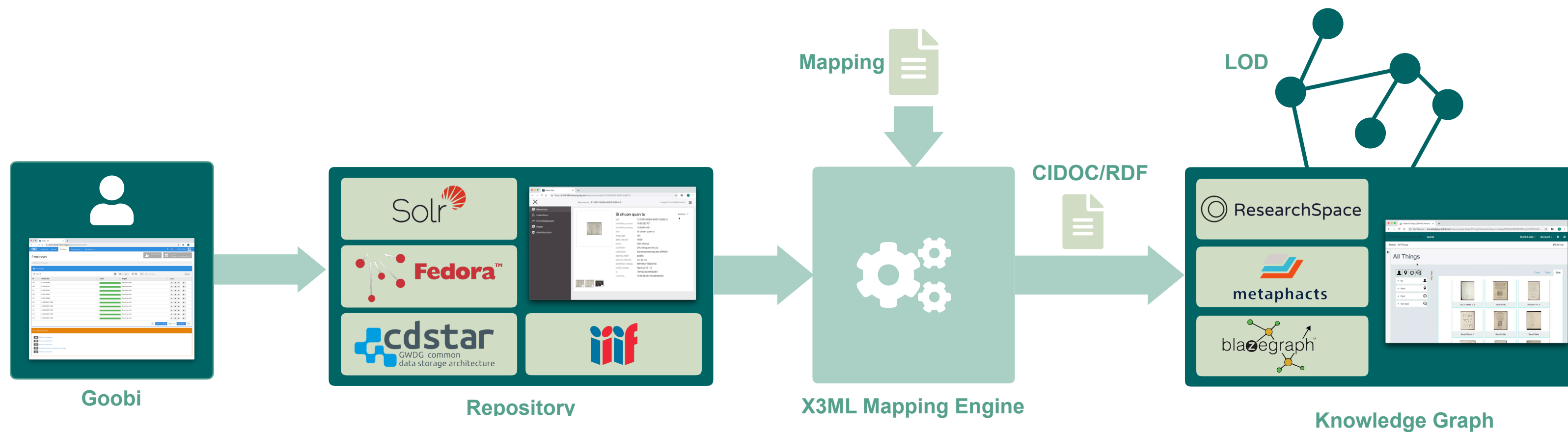


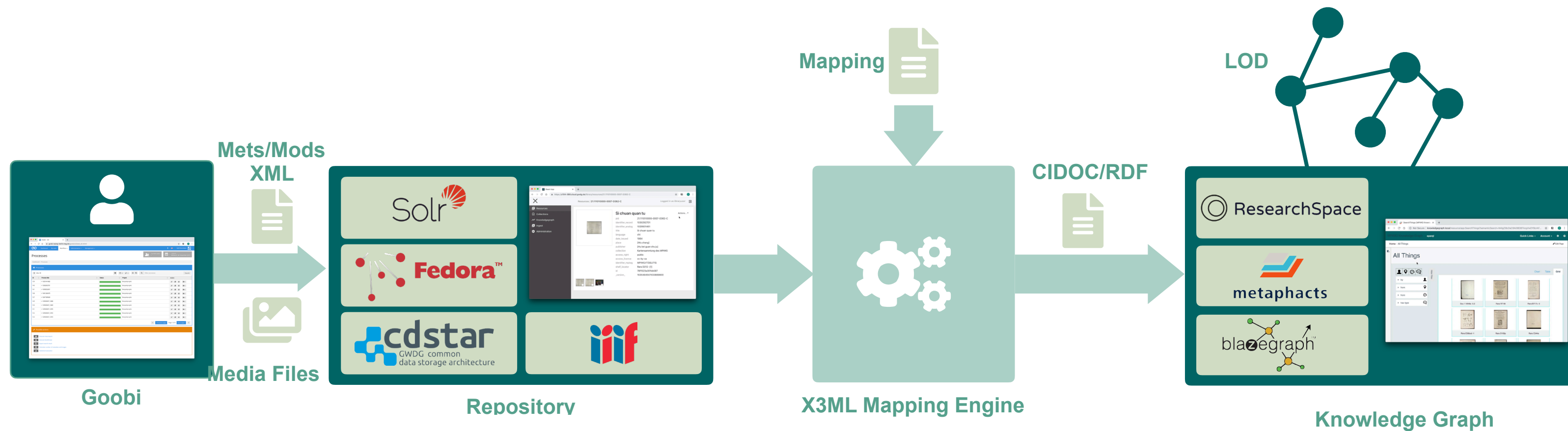
metaphacts



Knowledge Graph









Workflow

Research Data

[illegible]

Research Dataset



Workflow

Research Data

[illegible]

Research Dataset

Source Files



XML Version

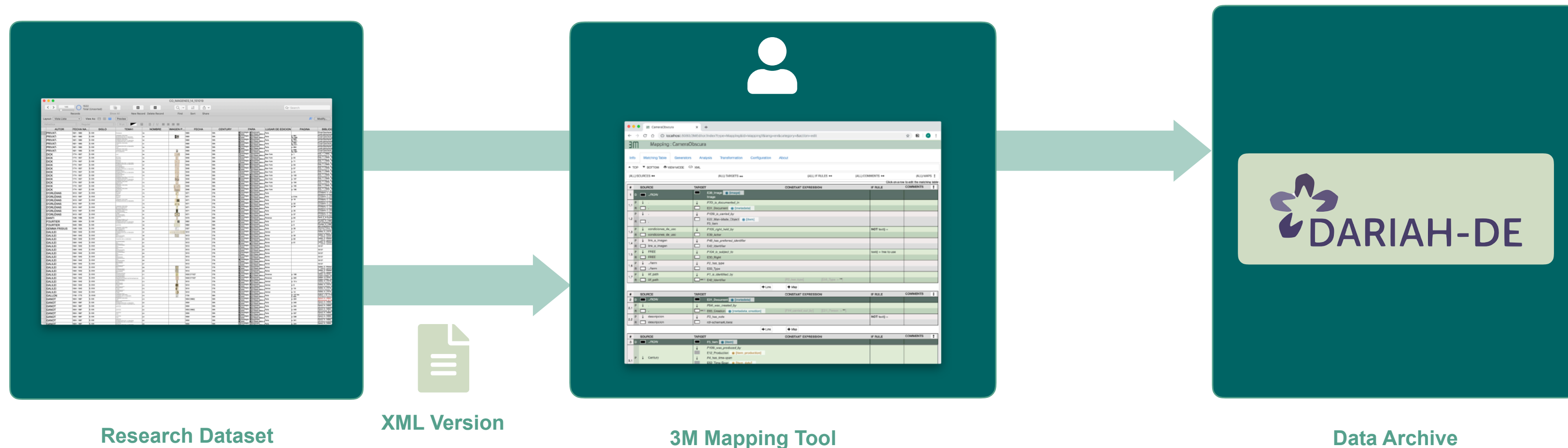
Research Dataset

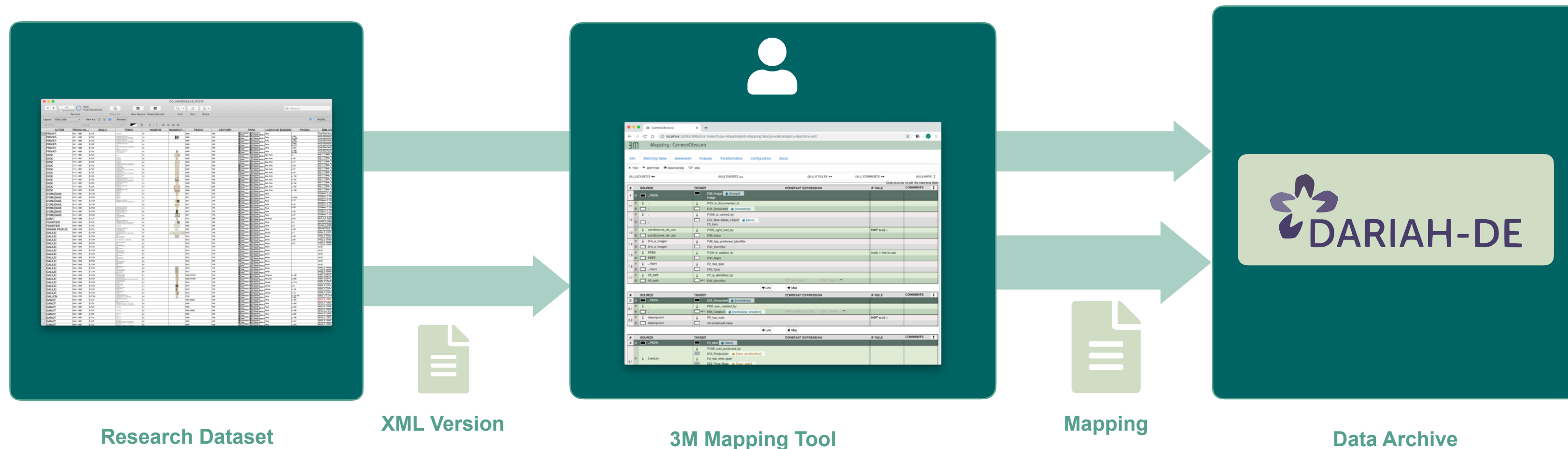
XML Version

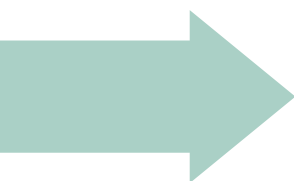


Data Archive



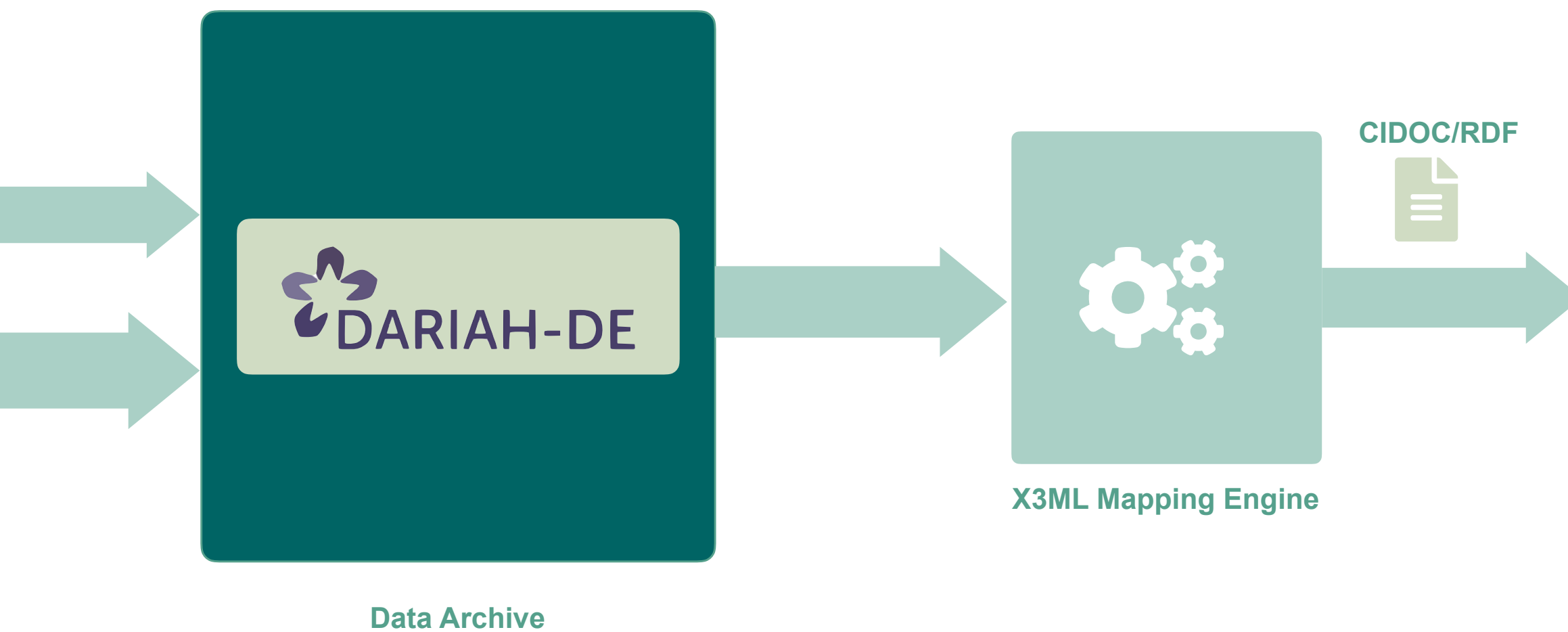


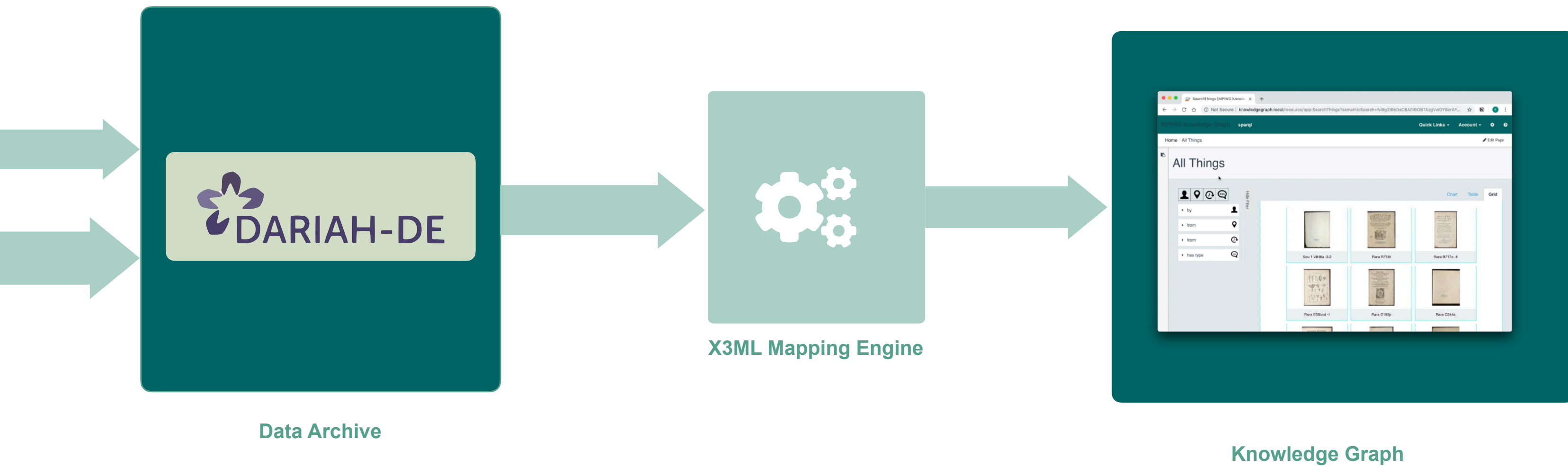


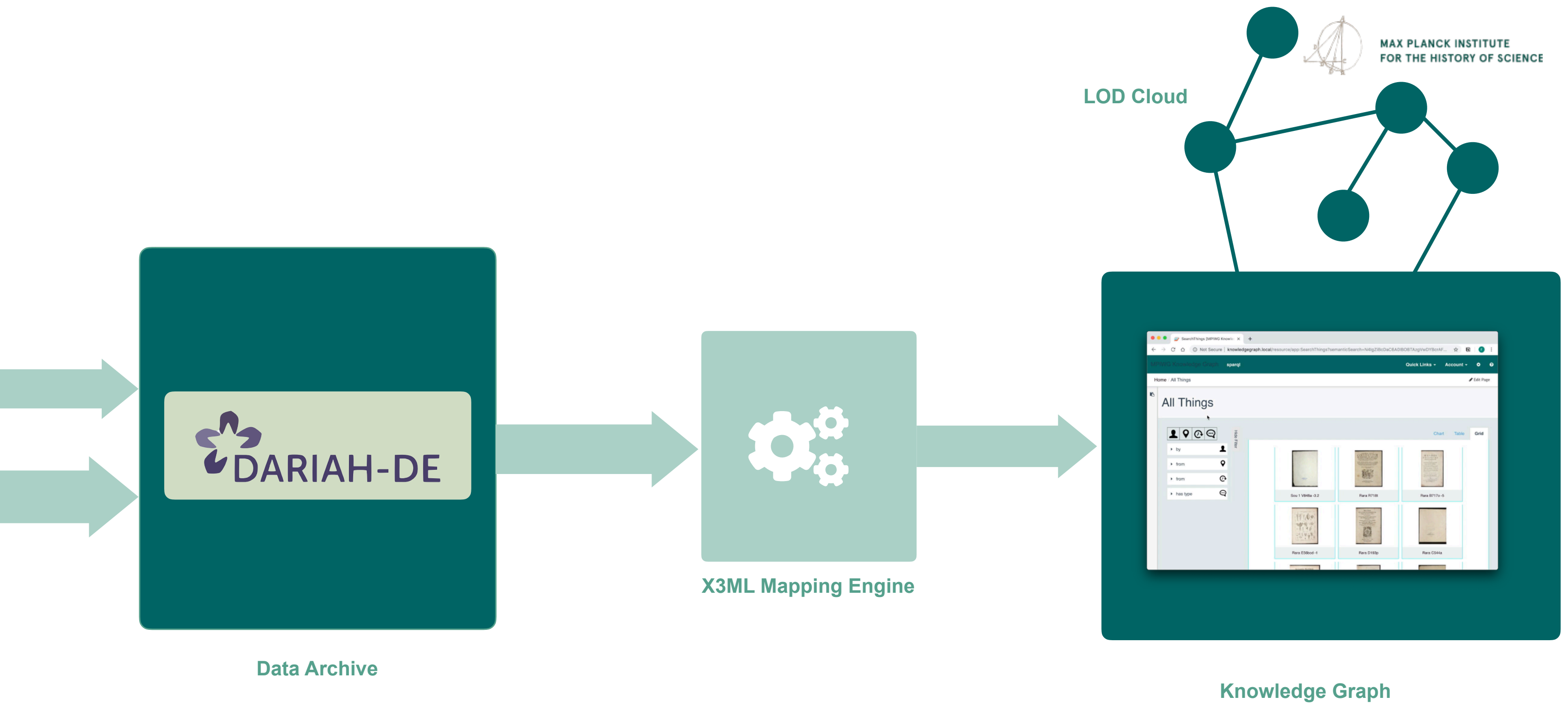


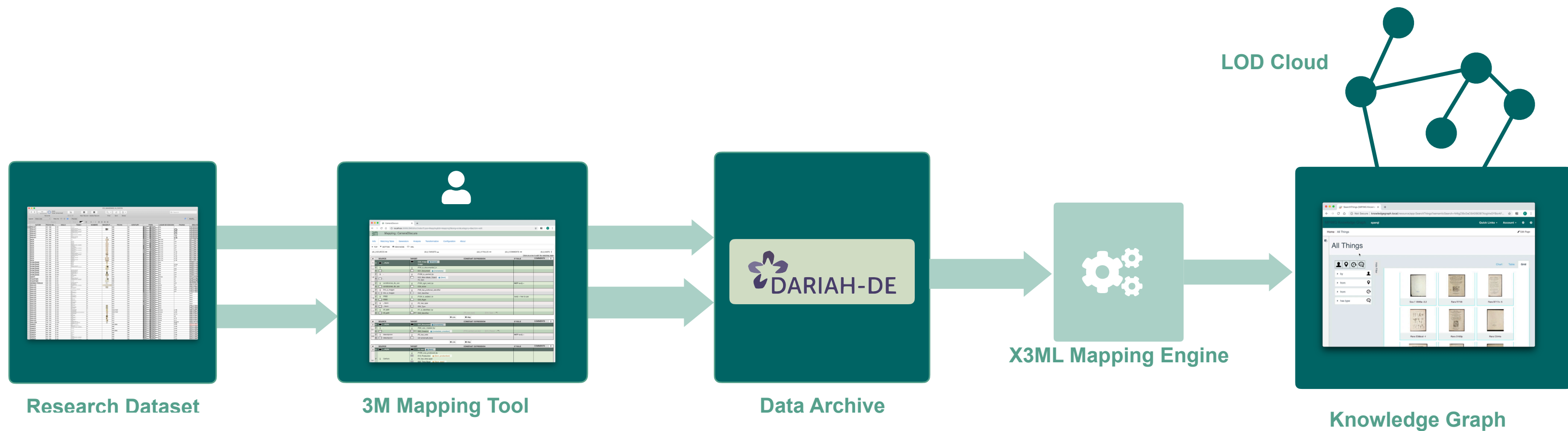
DARIAH-DE

Data Archive



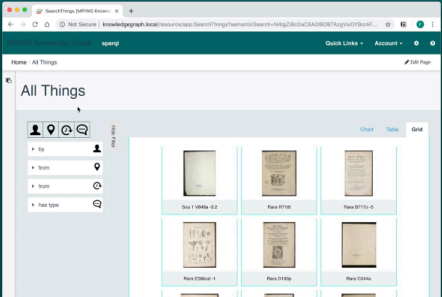








MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE

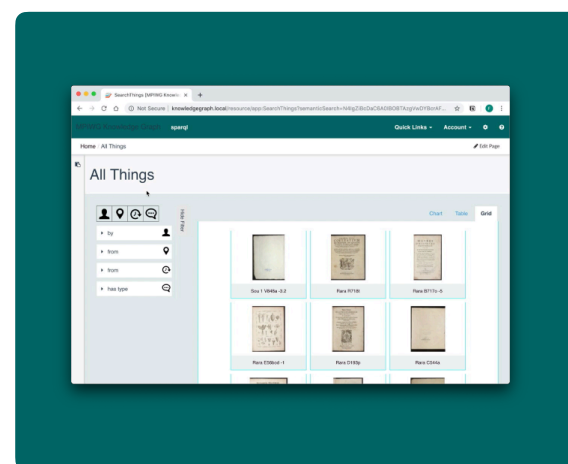


Knowledge Graph

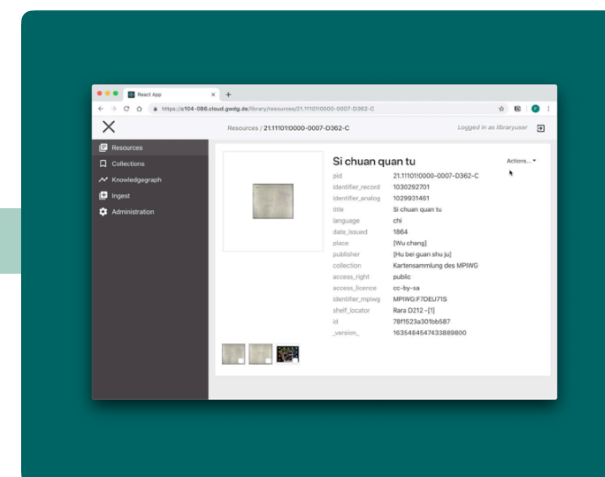


id	title	author	year	language	type	format	size	status	location	access	rights	notes
1	Si chuan quan tu	[Hu bei guan shu jù]	1984	zh	map	PDF	10MB	available	MPHG	public	CC-BY-SA	
2	Si chuan quan tu	[Hu bei guan shu jù]	1984	zh	map	PDF	10MB	available	MPHG	public	CC-BY-SA	
3	Si chuan quan tu	[Hu bei guan shu jù]	1984	zh	map	PDF	10MB	available	MPHG	public	CC-BY-SA	

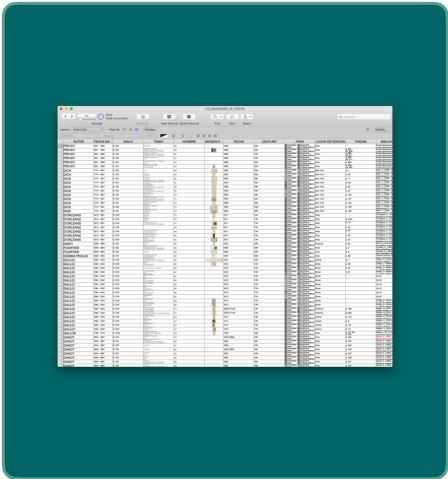
Research Dataset



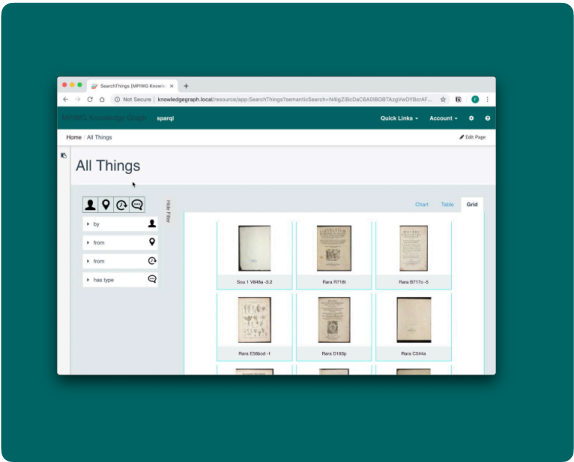
Knowledge Graph



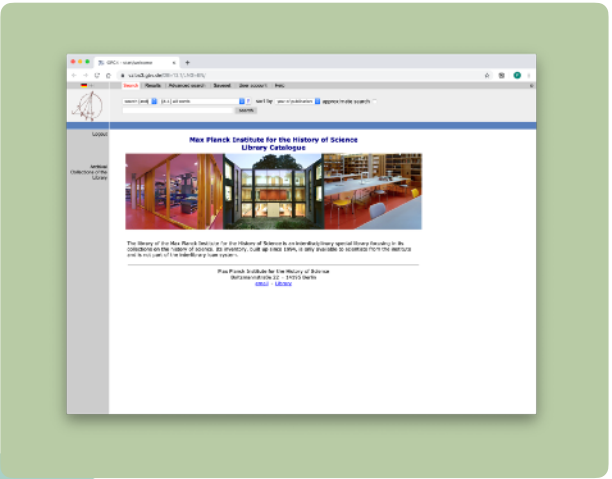
Repository



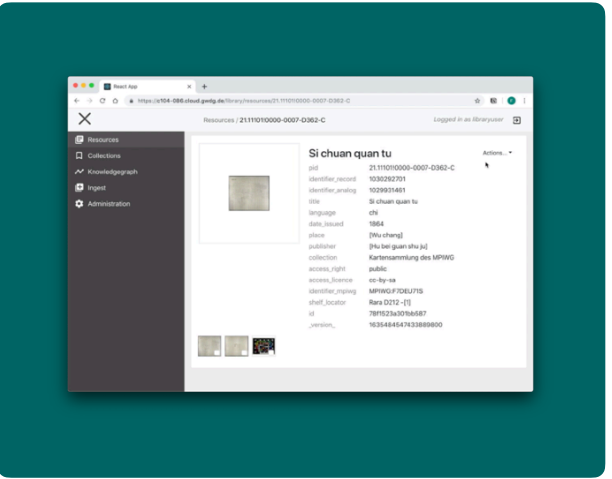
Research Dataset



Knowledge Graph



Library Catalogue



Repository



Publications

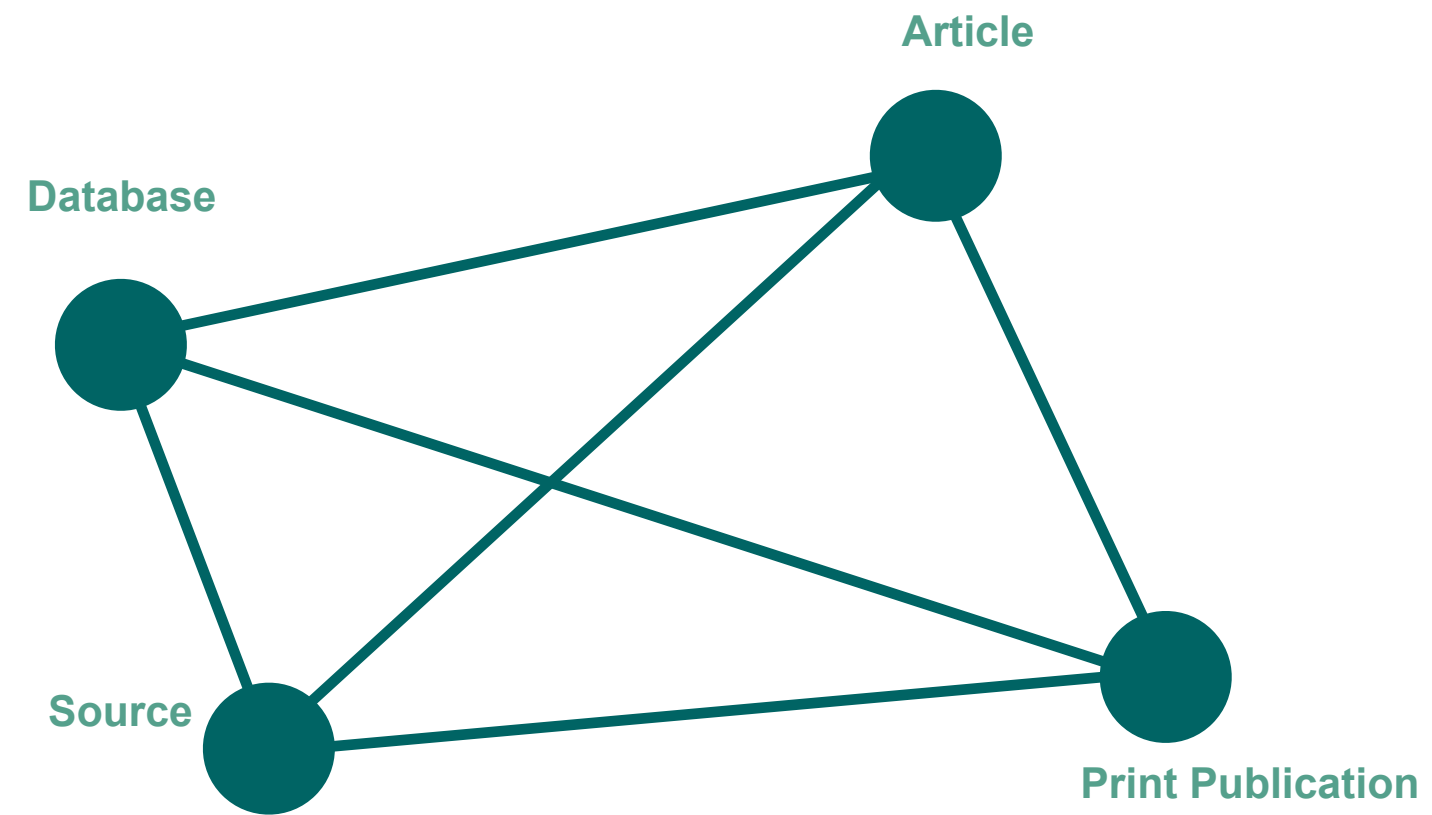


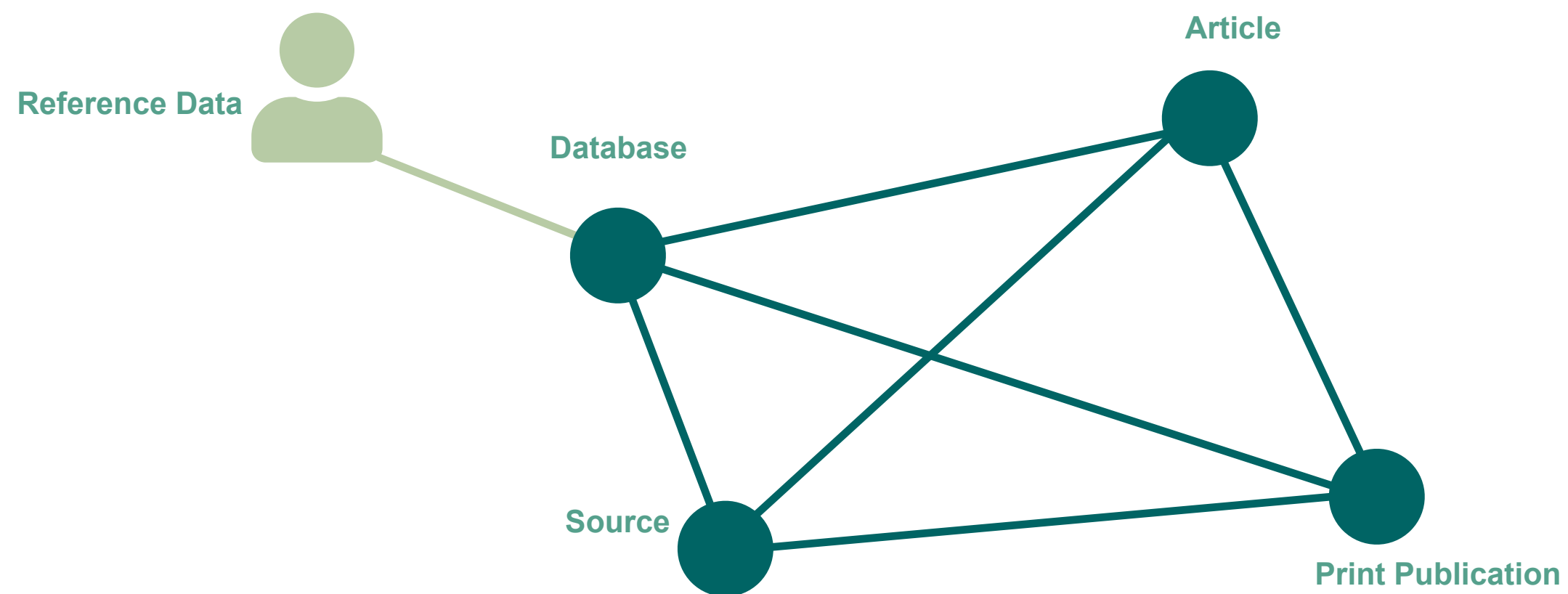
Database

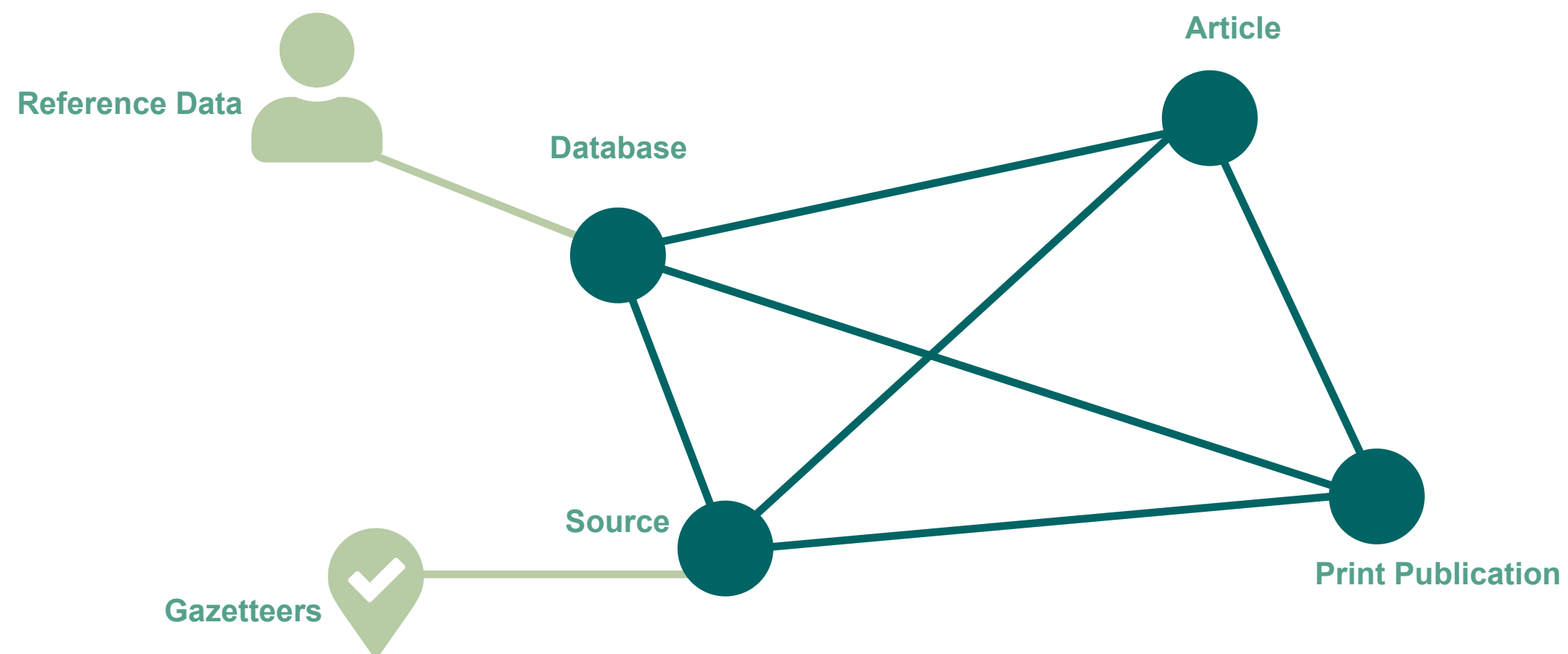


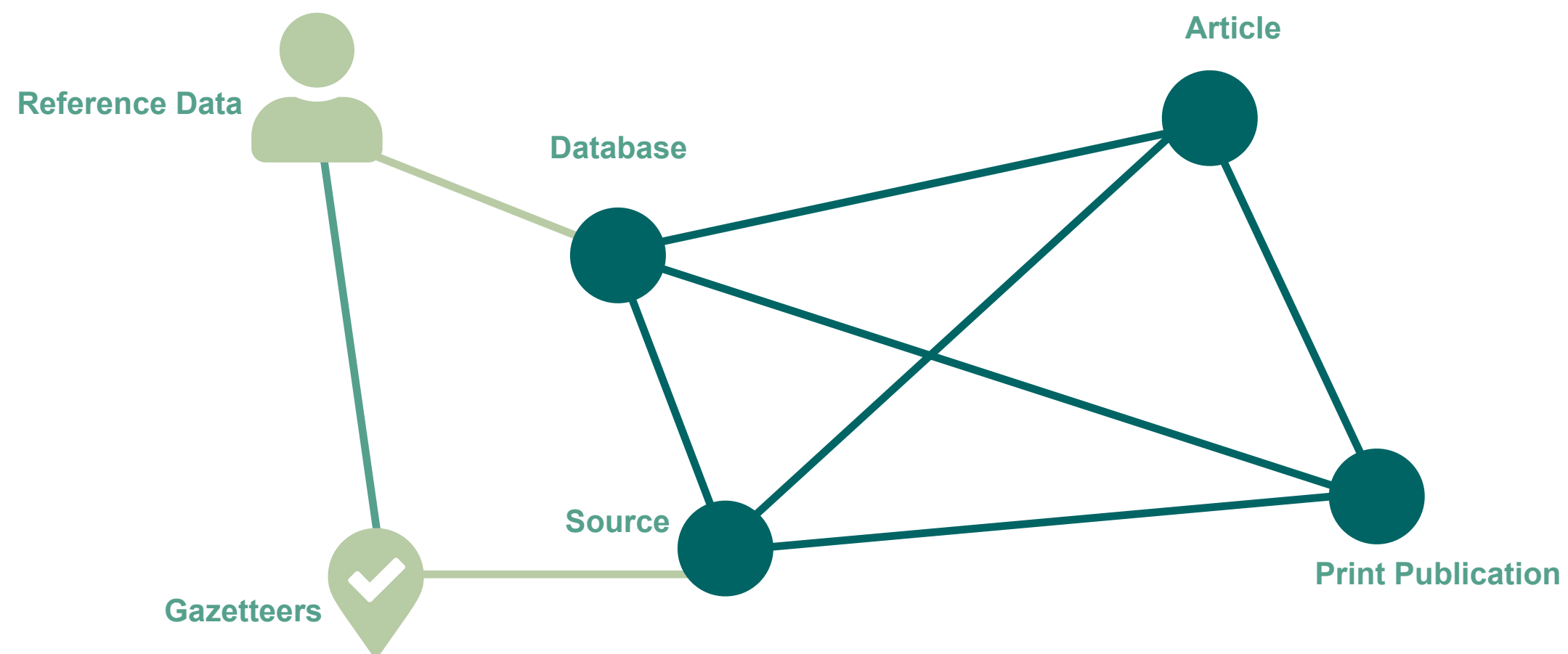
Source

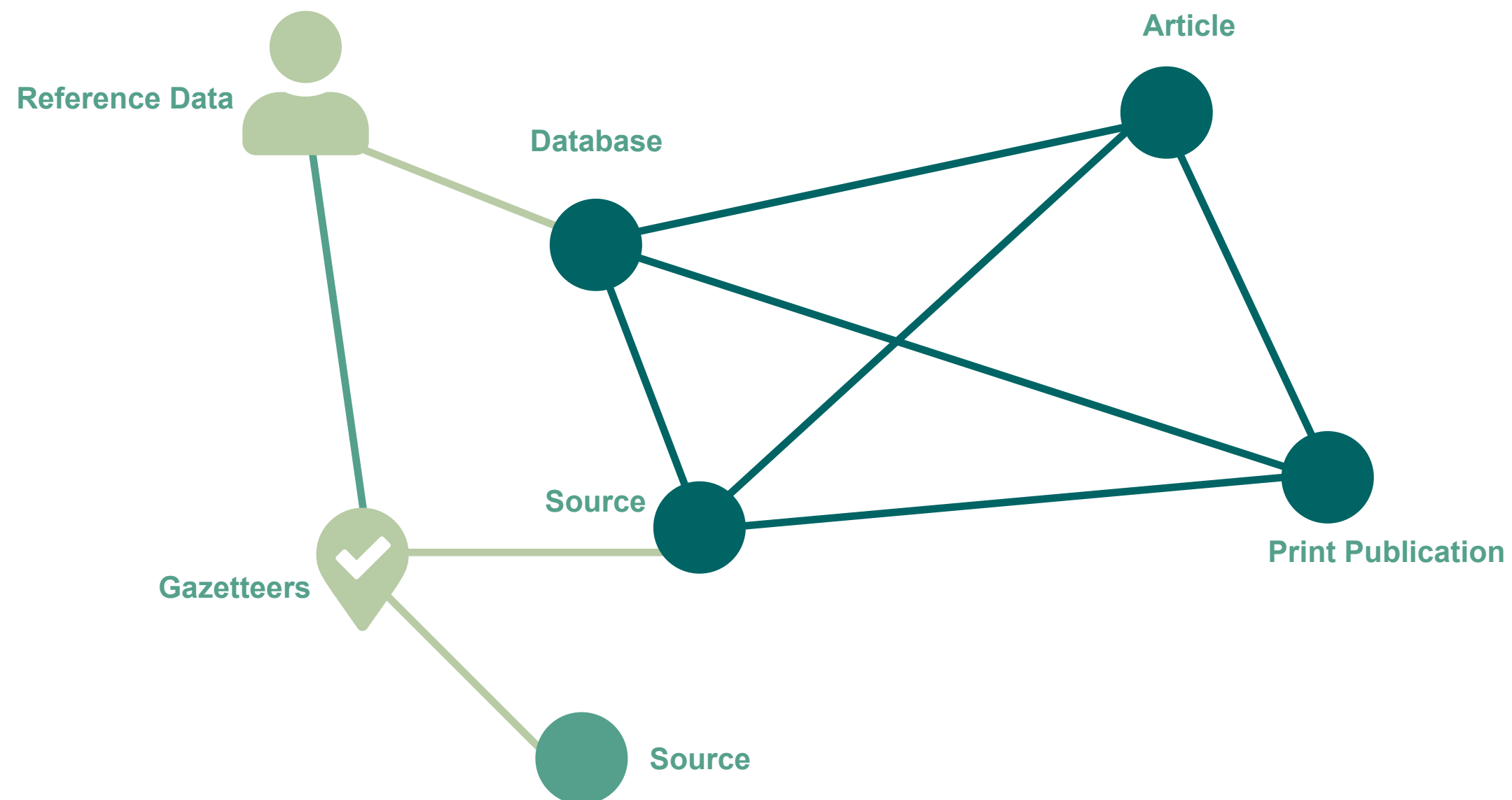


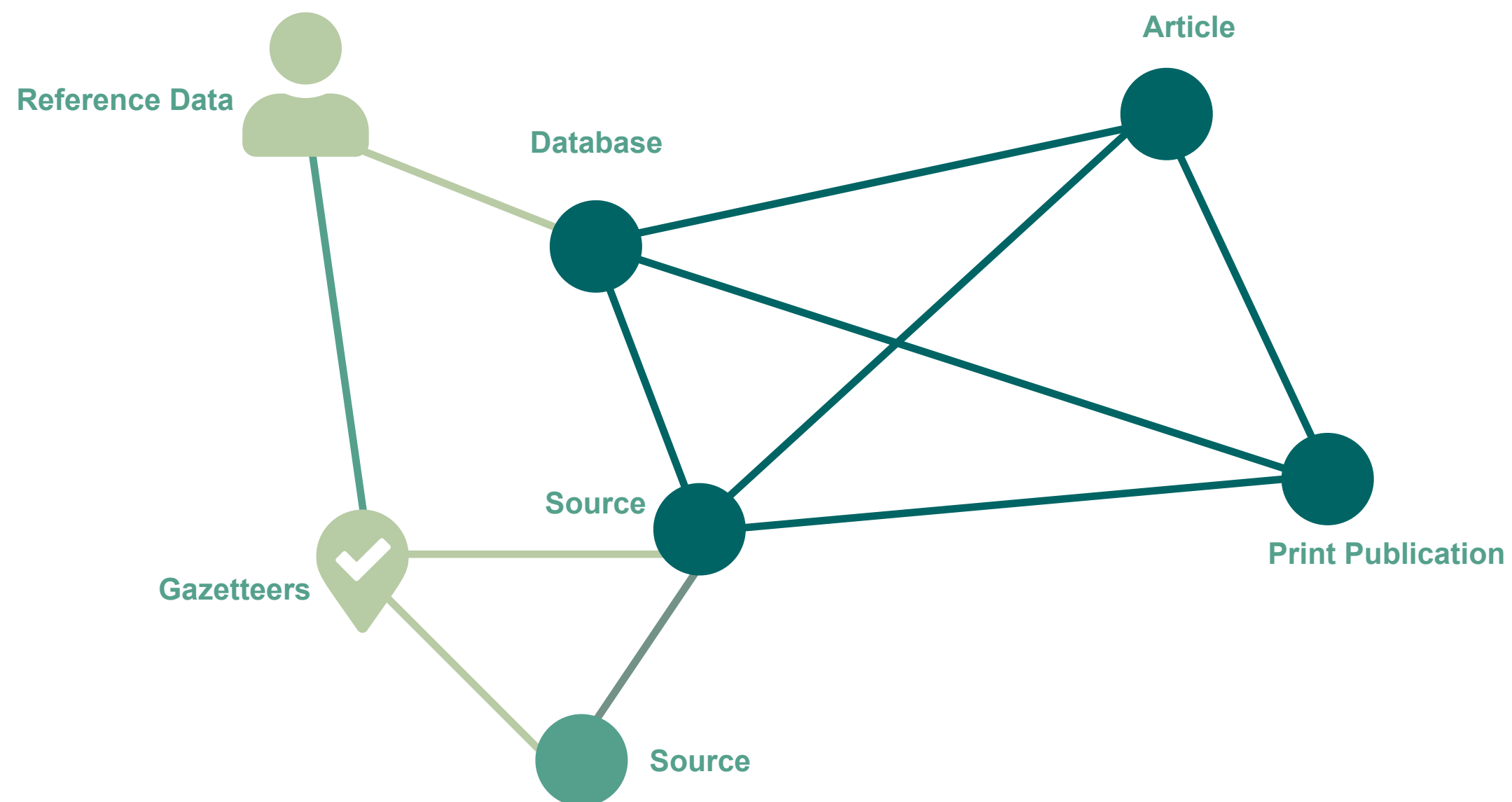














All Things

by

from

from

has type

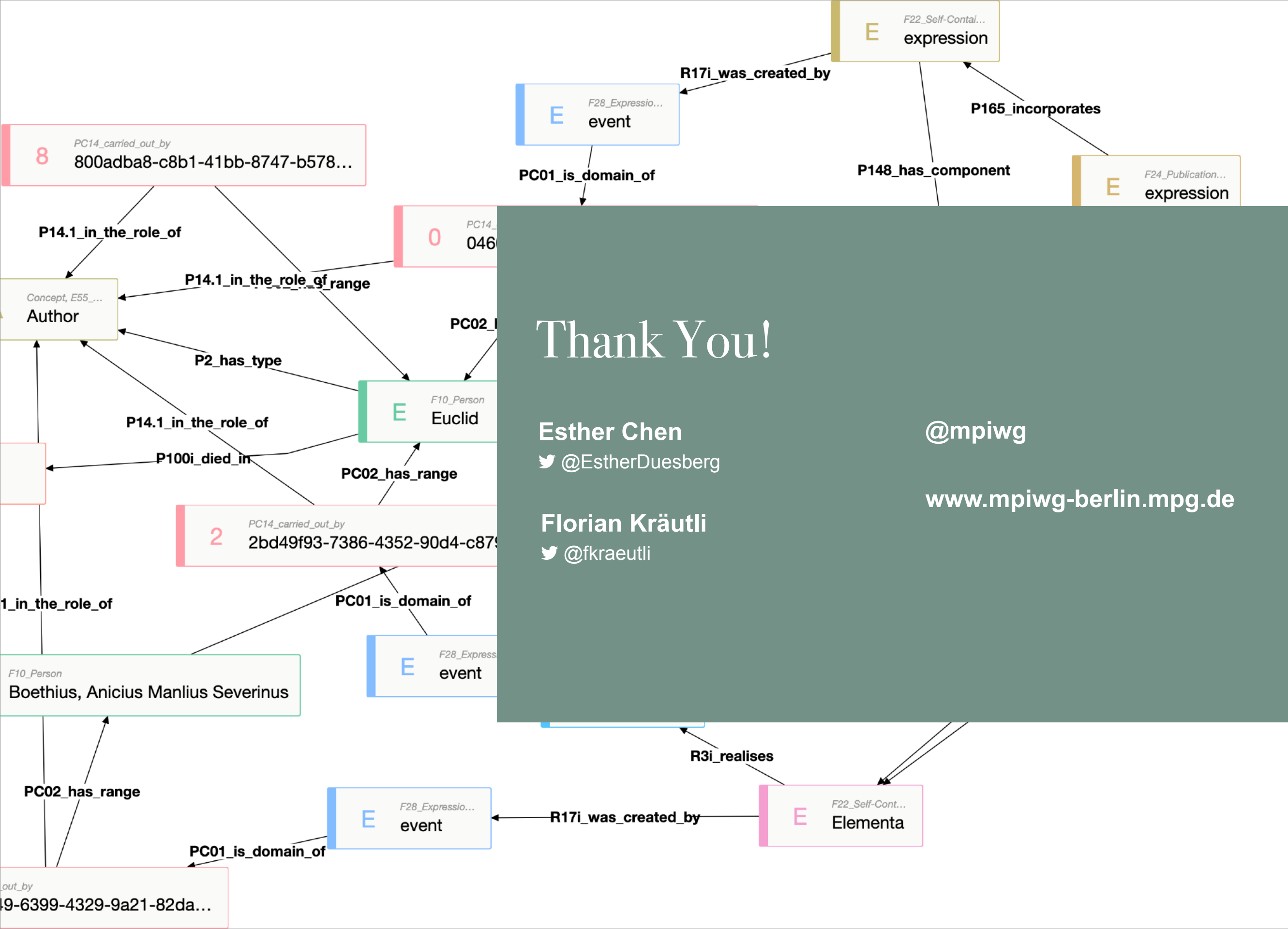
Hide Filter

Chart

Table

Grid

Sou 1 V848a -3.2	Rara R718t	Rara B717o -5
Rara E56bod -1	Rara D193p	Rara C544a



Thank You!

Esther Chen

🐦 @EstherDuesberg

@mpiwg

Florian Kräutli

🐦 @fkraeutli

www.mpiwg-berlin.mpg.de