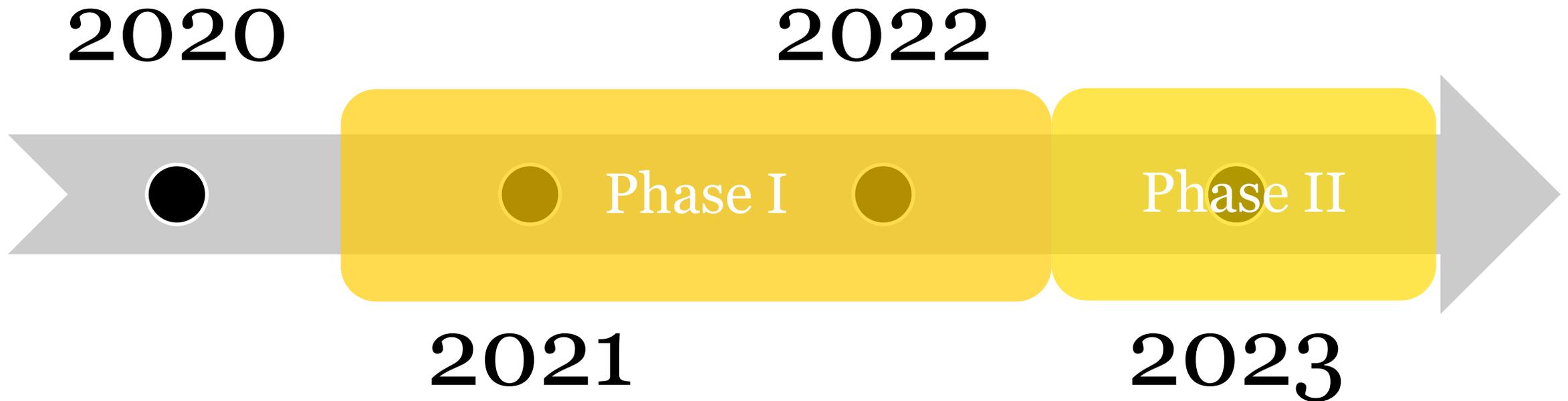


Künstliche Intelligenz im Landesarchiv Baden-Württemberg

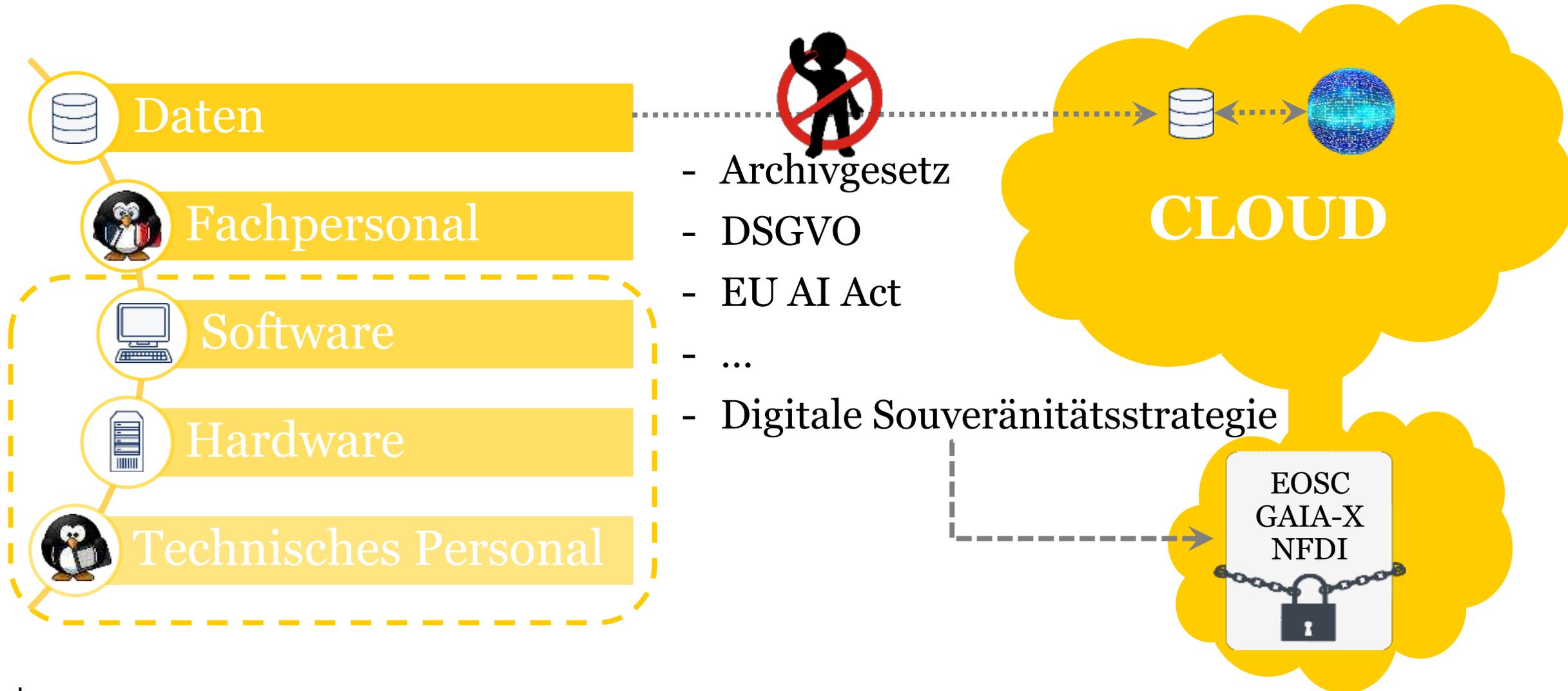
Deutscher Archivtag 2024 am 10.10.2024
Benjamin Rosemann

Laufzeit des FDMLab@LABW



Förderung durch die **Baden-Württemberg Stiftung** im Rahmen der **Zukunftsoffensive III**.

5 Zutaten für KI- und Daten-Projekte



Natural
Language
Processing

Knowledge
Representation

Automated
Reasoning

Machine
Learning

Computer
Vision

Robotic

Quellen:

- Stuart Russell und Peter Norvig, "What is AI?", in *Artificial Intelligence: A Modern Approach* (Upper Saddle River, N.J.: Prentice Hall, 2010), 2-3.

Clustering (Unüberwachtes Lernen)

OpenRefine



<https://openrefine.org/>

Name

Benjamin Rosemann
Benjamin Rosemann
Rosemann Benjamin
Beniamjn Rosemann
Benjamin Roseman
Benjamin Rosemam
Benjamin R. Rosemann

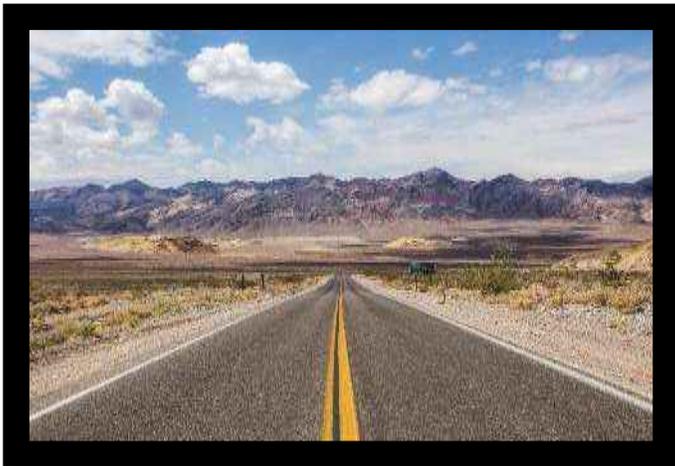
Orte

Nordschwaben (Rheinfelden)
Nordschwaben bei Rheinfelden
Nordschwaben-Rheinfelden
Rheinfelden-Nordschwaben
Nordschwaben / Rheinfelden
Rheinfelden Nordschwaben
Rheinfelden, Nordschwaben

Überwachtes Lernen



Training



6

Deutscher Archivtag 2024
Benjamin Rosemann



"Aptera 3"
von [Chris Anthony, CEO, Aptera Motors Corp.](#)
ist lizenziert unter einer [CC BY-SA 4.0](#) Lizenz.



"[Carl Benz Patent-Motorwagen Nummer 1](#)"
von [DaimlerChrysler AG](#)
ist lizenziert unter einer [CC BY-SA 3.0](#) Lizenz.

Named-entity recognition: Englisch

The State Archives **ORG** of Baden-Württemberg (**LABW ORG**) encompasses all state archives in **Baden-Württemberg GPE** .

The State Archives **ORG** of Baden-Württemberg is part of the portfolio of **the Baden-Württemberg Ministry for Science, Research and Arts FAC** .

It keeps historic documents that reach back to **the middle ages DATE** .

The current President of **the State Archives ORG** of Baden-Württemberg has been Prof. Dr. **Gerald Maier PERSON** since **February 1st, 2018 DATE** .

It employs **around 250 CARDINAL** people at **8 CARDINAL** locations.

Das Model *en_core_web_lg* hat einen **f-score** von **0,85** und unterscheidet **18** Arten von **Entitäten**:
CARDINAL, DATE, EVENT, FAC, GPE, LANGUAGE, LAW, LOC, MONEY, NORP, ORDINAL, ORG,
PERCENT, PERSON, PRODUCT, QUANTITY, TIME und WORK_OF_ART.

Quellen:

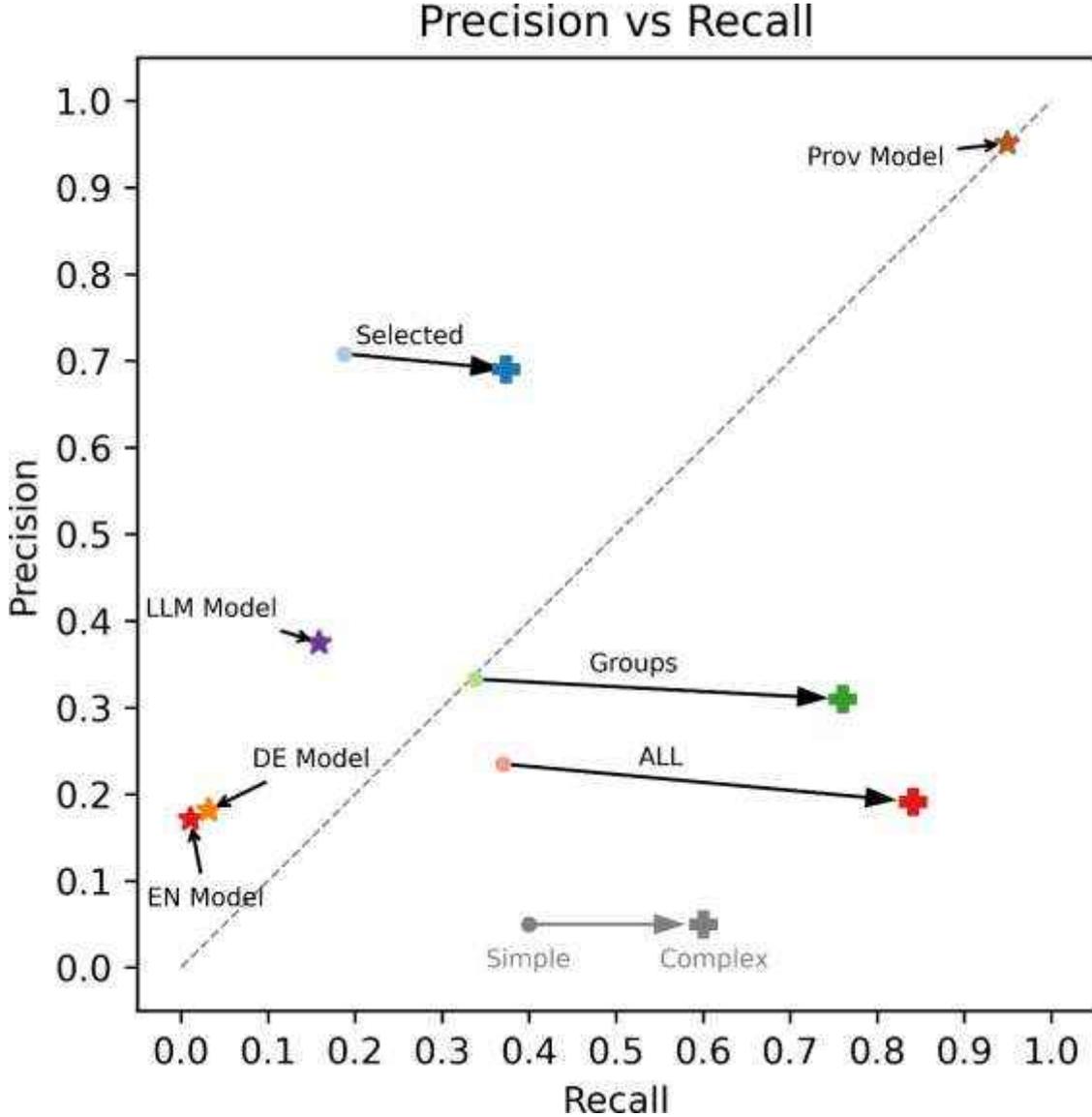
- Erstellt mit spaCy 3.3 *en_core_web_lg* (https://spacy.io/models/en#en_core_web_lg).
- Wikipedia Autoren, „Landesarchiv Baden-Württemberg“, Wikipedia – Die freie Enzyklopädie, https://de.wikipedia.org/w/index.php?title=Landesarchiv_Baden-W%C3%BCrttemberg&oldid=223510351 (aufgerufen am 20. Juli 2022).

Large Language Modelle



Was ist die GND-ID von Mozart?

Die GND Nummer von Wolfgang Amadeus Mozart in der Gemeinsamen Normdatei lautet "119331550".



NFDI4Culture: Antelope



Antelope is a powerful free, open source service for easy data annotation: terminology search; entity linking; and integration with research data management systems.

TERMINOLOGY SEARCH ENTITY LINKING IMAGE RECOGNITION

Type the term you want to search for (single word or phrase): ⓘ

Vincent van Gogh

SELECT TERMINOLOGY SOURCES ⓘ

- WIKIDATA
- WIKIDATA + DBpedia
- ICONCLASS
- GND (Gemeinsame Normdatei)
- TIB Terminology Service

SETTINGS

- Allow class duplicates ⓘ

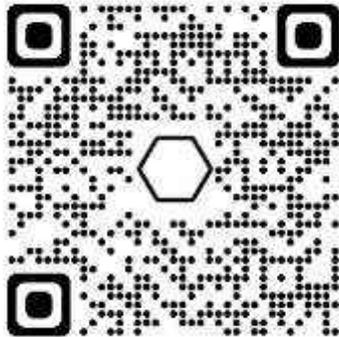
SEARCH **RESET**



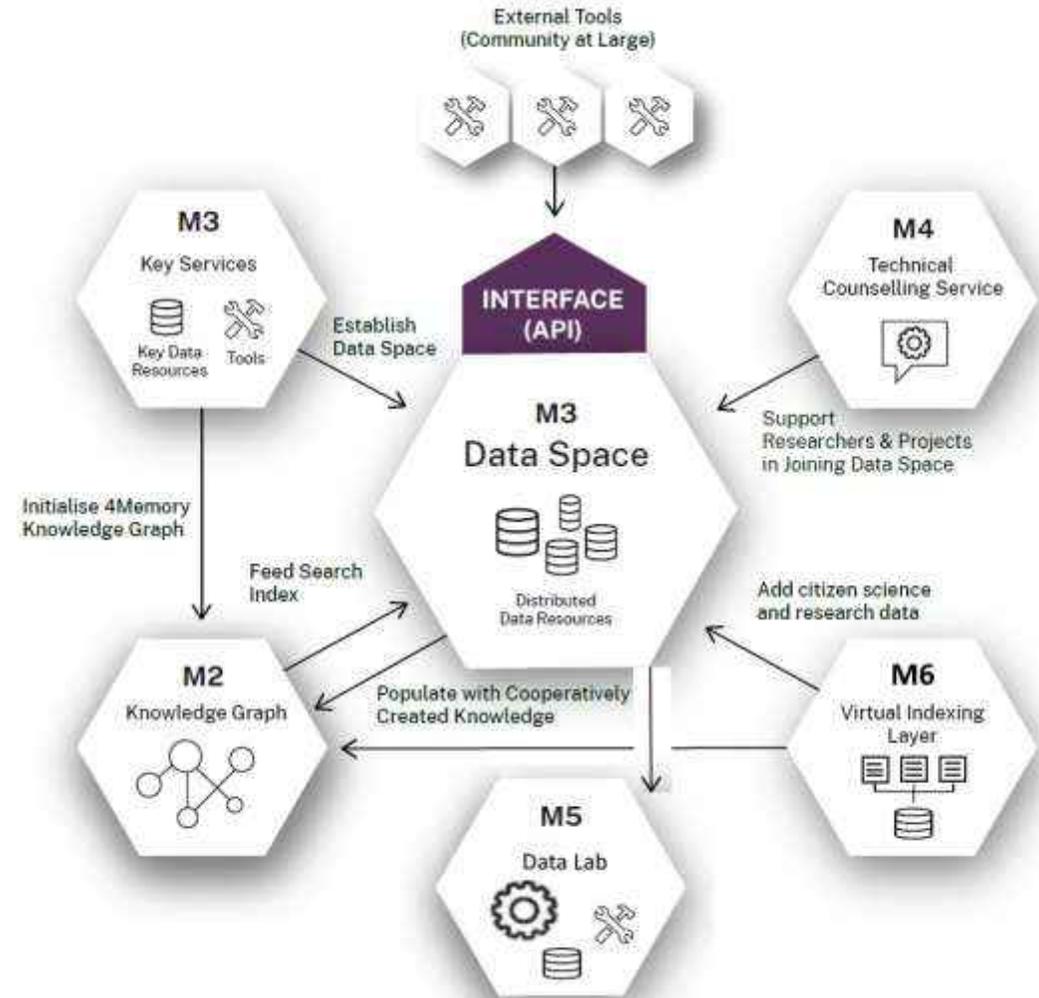
<https://service.tib.eu/annotation/>

Task Area 3: Data Services

- **M5: Data Lab**
- High Performance Cluster
- Passende (online) Werkzeuge



<https://4memory.de/>



Quellen:

- Daniel Fähle, Harald Sack, "Perspektive „digitaler Werkzeugkasten“ für historische Forschung mit Archivgut" (Präsentation), 53. Deutscher Historikertag, München, 06. Oktober 2021.

Vielen Dank!

Benjamin Rosemann

Landesarchiv Baden-Württemberg
Zentrale Dienste

URL: <https://www.landearchiv-bw.de>

Tel.: +49 711 335075-512

E-Mail: benjamin.rosemann@la-bw.de

 <https://orcid.org/0000-0002-0780-3979>

