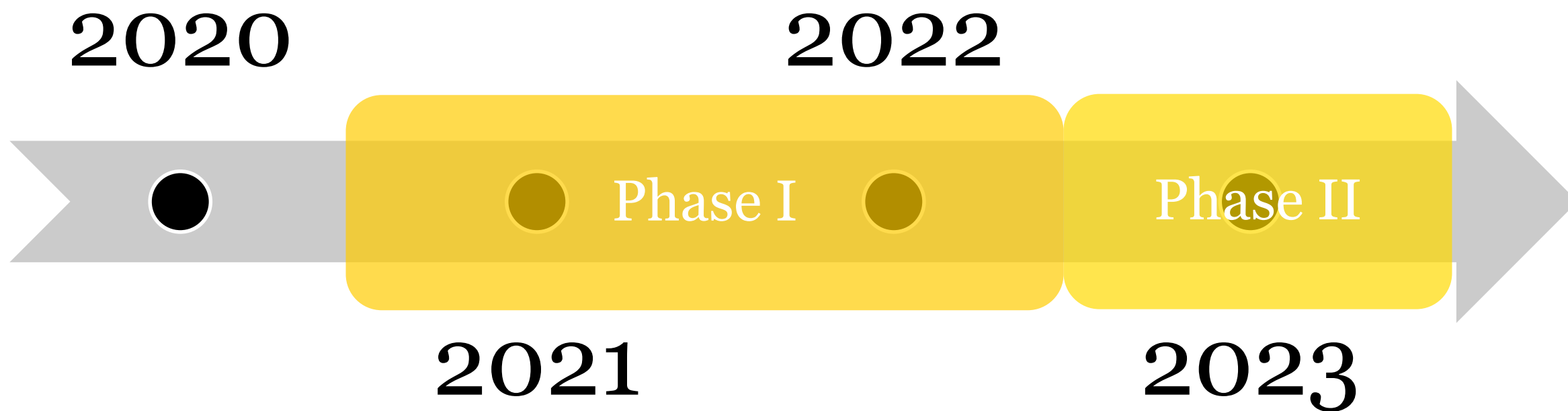


Künstliche Intelligenz im Landesarchiv Baden-Württemberg

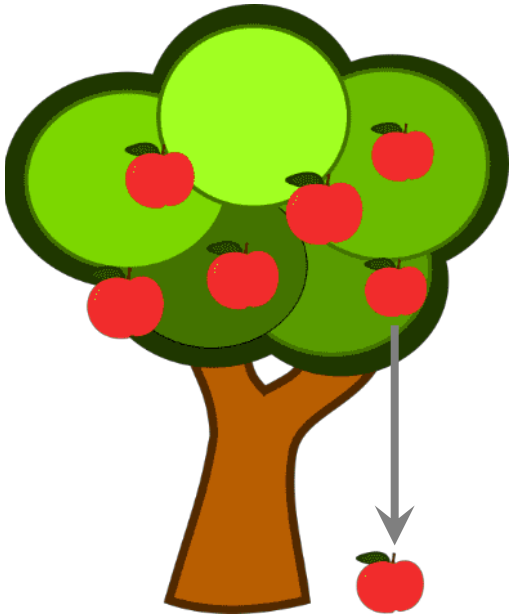
Schleswig-Holsteinischer Archivtag 2024 am 24.04.2024
Benjamin Rosemann



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

Verständlich und exakt?

Wahrnehmung >



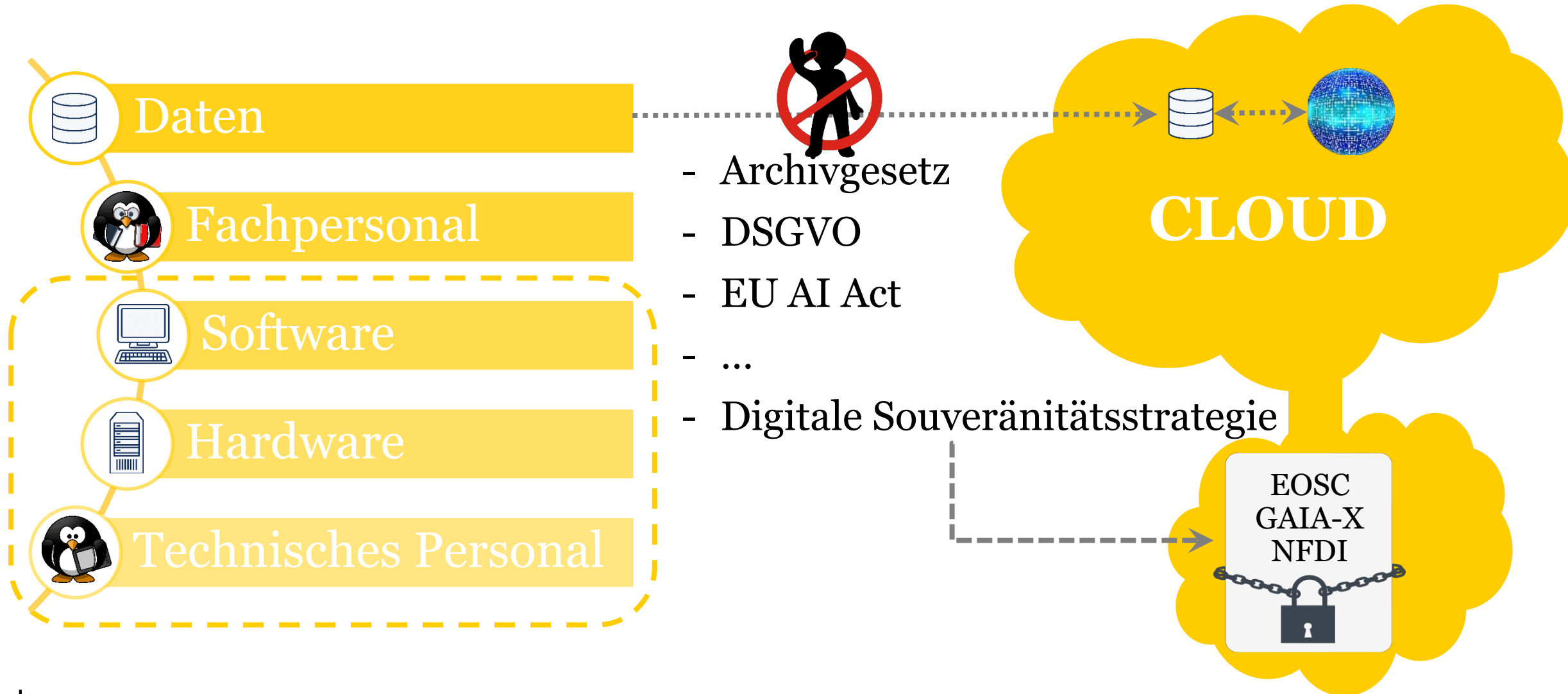
Mechanik >

$$\vec{G} = m \cdot g$$

Quantenmechanik

...

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.

Die Rechenmaschine



Es ist Sonntag Nachmittag und es regnet. Laut Werbeprospekt kosten 1kg Äpfel im Supermarkt 2 Euro.

Ich möchte 5kg Äpfel. Wie viel Geld benötige ich zum Einkaufen?

Wie funktioniert das Training für ein ML Modell?



Training



"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.

Volltexte erstellen mit ML Unterstützung



OCR-D

<https://ocr-d.de/>



Transkribus

<https://transkribus.eu/>



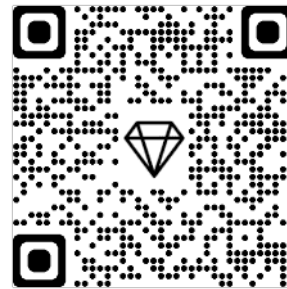
eScriptorium

<https://gitlab.com/scripta/escriptorium>



OpenRefine is like Excel on steroids!

Quelle unbekannt



<https://openrefine.org/>

OpenRefine Gurs

Facet / Filter

Undo / Redo 10 / 10

Refresh

Reset All

Remove All

Geschlecht (m/w)

change

17 choices Sort by: name count

Cluster

m 1758

maennlich 9

man 3

man. 10

manlich 8

mannlich 9

mannlich 4

mannlich 22

w 1973

weblich 1

wei. 2

Familienverhältnis

change

19 choices Sort by: name count

Cluster

alleinstehend 1

gesch. 1

geschieden 3

ledig 29

single 2

verehatet 1

verh. 6

verheirat 1

verheiratet 1575

verhieratet 1

verhiratet 1

Todesdatum (Jahr)

change reset

1,940 — 2,022

☒ Numeric

☐ Non-numeric

☐ Blank

☐ Error

3936

0

14

0

3986 matching rows (4000 total)

Extensions: Wikidata

Show as: rows records

Show: 1 5 10 25 50 100 250 500 1000 rows

« first

« previous

3 of 160 pages

next »

last »

		Geschlecht (m/w)	akad. Grad/Titel	Familienverhältnis	Geburtsdatum (exakt)	Geburtsort	Geburtsort GND-ID	Religion	Todesdatum (Tag)	Todesdatum (Monat)	Todesdatum (Jahr)	Todesdatum (beschreibend)
☆	🗨	51. w			16.09.1869	Sankt Goar	4051602-7	israelitisch	1	9	1943	ermordet
☆	🗨	52. m			26.12.1878	Tauberbischofsheim	4059128-1	israelitisch	19	12	1943	verschollen
☆	🗨	53. m			31.03.1863	Schopfloch	4204796-1	israelitisch	8	8	1945	ermordet
☆	🗨	54. w			09.08.1855	Altdorf bei Ettenheim	4260757-7	israelitisch	8	5	1958	für tot erklärt, ermordet
☆	🗨	55. w			19.06.1892	Worms	4068942-7	israelitisch	19	8	1941	verschollen
☆	🗨	56. w		verheiratet	25-09-1885	Lichtenau	4690984-9	israelitisch	18	1	1955	
☆	🗨	57. m		verheiratet	30-03-1882	Schwegenheim	4235625-8	israelitisch	8	2	1941	
☆	🗨	58. m		verwitwet	16. August 1867	Stuttgart	4058282-6	 Stuttgart (4058282-6)		2	1945	
☆	🗨	59. m		verheiratet	15.07.1879	Königsbach	4132021-9			5	1945	ermordet
☆	🗨	60. m			29.11.1879	Karlsruhe	4029713-5			5	1940	ermordet
☆	🗨	61. m			30-05-1868	Schriesheim	4053296-3			11	1948	ermordet
☆	🗨	62. m		verheiratet	01.08.1884	Gemünden im Hunsrück	2001763-7	israelitisch	8	5	1942	verschollen
☆	🗨	63. w		verwitwet	12.12.1876	Berwangen	4280920-4	israelitisch	8	10	1941	ermordet
☆	🗨	64. m			06-08-1874	Richen	7830878-8	israelitisch	31	12	1942	ermordet
☆	🗨	65. w			15.10.1877	Sierck-les-Bains (Frankreich)	4054905-7	israelitisch	8	8	1945	ermordet
☆	🗨	66. w			01.08.1921	Tauberbischofsheim	4059128-1	israelitisch	8	5	1945	
☆	🗨	67. w		verehatet	08.04.1872	Dreibrunnen (Deutsches Reich heute Frankreich)	4817592-4	israelitisch	5	6	1945	für tot erklärt, ermordet
☆	🗨	68. w		verheiratet	12-03-1878	Kiel	4030481-4	israelitisch	22	1	1945	
☆	🗨	69. m			16-03-1872	Lichtenau	4690984-9	israelitisch	8	3	1942	verschollen
☆	🗨	70. m			05.11.1887	Bonfeld	4593548-8	israelitisch	8	5	1941	für tot erklärt, ermordet
☆	🗨	71. w		verheiratet	06-06-1879	Eschenau	4509954-6	israelitisch	28	9	1962	für tot erklärt, ermordet
☆	🗨	72. m			09-04-1867	Bendzin bei Pless (Russisches Kaiserreich heute Polen)	4559259-7	israelitisch	8	8	1945	für tot erklärt, ermordet
☆	🗨	73. m		verheiratet	19.10.1896	Reichenbach (Lautertal-Reichenbach)	7564415-0	israelitisch	8	1	1945	verschollen
☆	🗨	74. m			29.06.1878	Sinsheim	4055123-4	israelitisch	21	10	1956	für tot erklärt, ermordet
☆	🗨	75. m	Dr. Med.		26.04.1887	Bad Neustadt a. d. Saale	4041939-3	israelitisch	8	9	1952	für tot erklärt, ermordet

10

Schleswig-Holsteinischer Archivtag 2024
Benjamin Rosemann

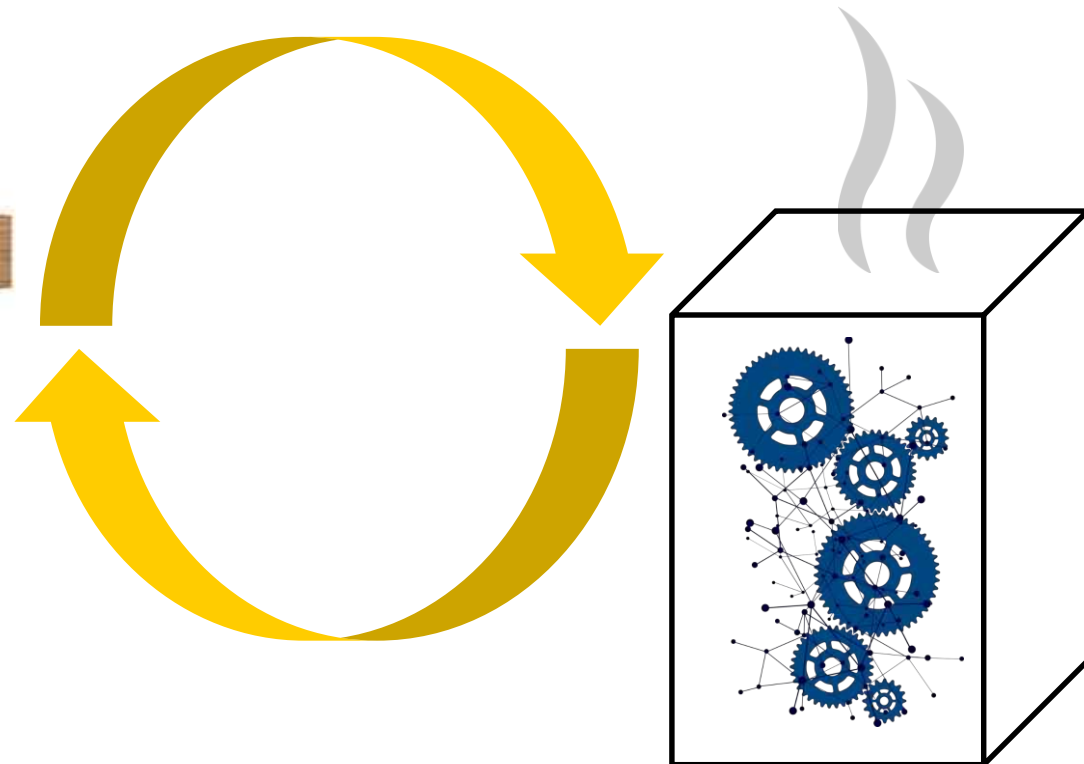
Integrierte Algorithmen

- FP: Fingerprinting
- FP: N-Gram Fingerprinting
- Phon: Cologne-phonetic
- Phon: Methaphone3
- Phon: Daitsch-Moktoff
- Phon: Baider-Morse
- NN: Levensthein
- NN: Prediction by Partial Matching

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

Human-in-the-loop



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).

Johannes Stoffel , Museumsdirektor der kurfürstlichen Museen in Augsburg



Frau Karolina Löb , geb. Reis , geboren am 04.03.1862 in Alsheim , gestorben in Gurs , Odenbach am Glan



Kaiser Friedrich I. , genannt Barbarossa





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/>

Objekterkennung auf Bildern?



Die Rechenmaschine



Es ist Sonntag Nachmittag und es regnet. Laut Werbeprospekt kosten 1kg Äpfel im Supermarkt 2 Euro.

Ich möchte 5kg Äpfel. Wie viel Geld benötige ich zum Einkaufen?

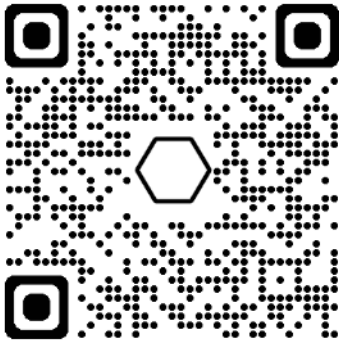
Die Suchmaschine?



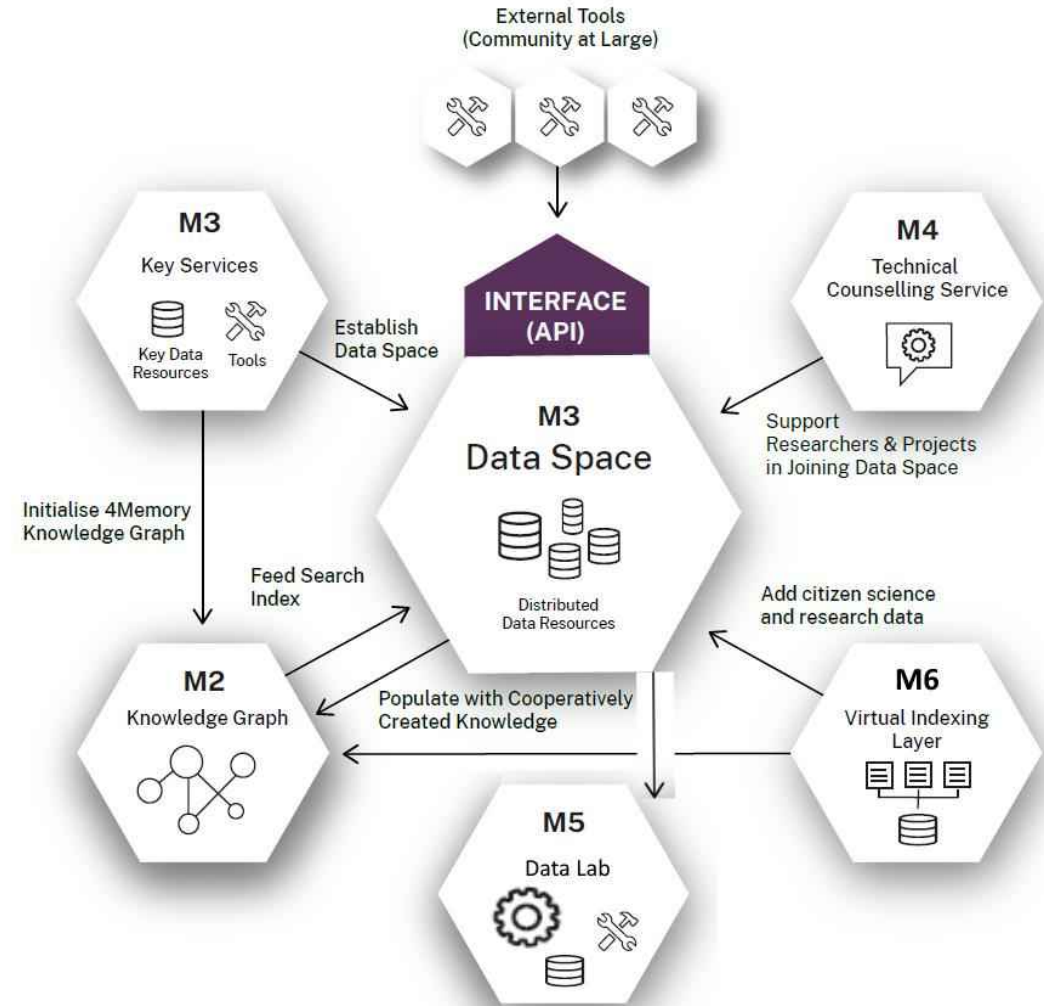
Was ist die GND-ID von Mozart?

Task Area 3: Data Services

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



<https://4memory.de/>



Quellen:

- Daniel Föhle, 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.landesarchiv-bw.de>

Tel.: +49 711 335075-512

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

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

