

Universität Stuttgart



DA-FDM – Ein Digitaler Assistent zur
halbautomatischen Erschliessung
von Forschungsdaten

Dr. Peter Schäuble, Dr. Helge Steenweg

2. November 2023

ubs

Digitaler Assistent für Forschungsdatenmanagement (DA-FDM) baut auf dem DA-3 auf



<https://www.da-3.eu>

Ein Projekt

- der UB Stuttgart,
- des BSZ und der
- Fa. Eurospider, Zürich

Seit 2016

DA-2 und DA-3
jeweils gefördert vom MWK im
Rahmen IBS:BW



Gefördert durch das MWK BW



DA3

The screenshot displays the GND interface with several red circles highlighting specific elements:

- 1:** Points to the search results list on the left, where 'Qualitative Sozialforschung' is selected.
- 2:** Points to the 'Bestand' (Content) field in the 'Info > GND' section, showing 's [Sachschließung]'.
- 3:** Points to the 'Publ.' (Publication) field in the 'Kurztitel' section, showing 'London : Routledge, 2021'.
- 4:** Points to the 'BK' (Bibliography Classification) field in the 'Einschließung' section, showing '76.20 Sportorganisation'.
- 5:** Points to the 'Vorschläge' (Suggestions) list in the 'Vorschläge' section, where 'Qualitative Sozialforschung [Sach]' is highlighted.
- 6:** Points to the 'Sport' category in the 'Einblendungen' (Display) section.



Verbreitung des Digitalen Assistenten für die halbautomatische Sacherschließung



<https://wiki.k10plus.de/display/K10PLUS/Digitaler+Assistent+DA-3>

<https://www.da-3.de>

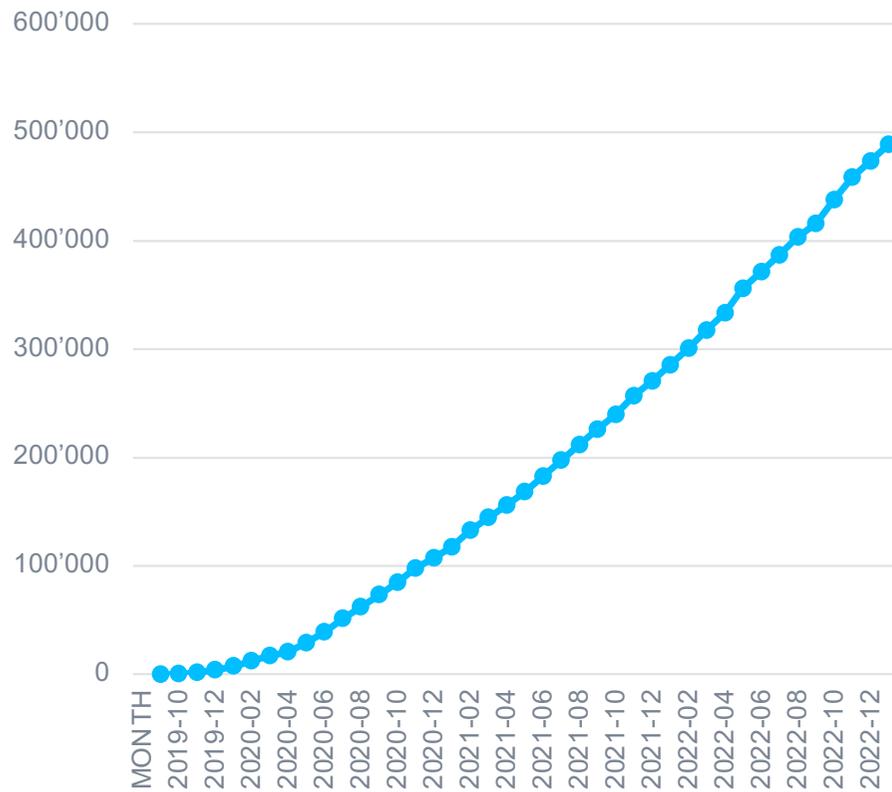
<https://www.bsz-bw.de/kooperationen/da-3.html>

<https://www.eurospider.com/de/relevancy-produkt/digitaler-assistent-da-3>

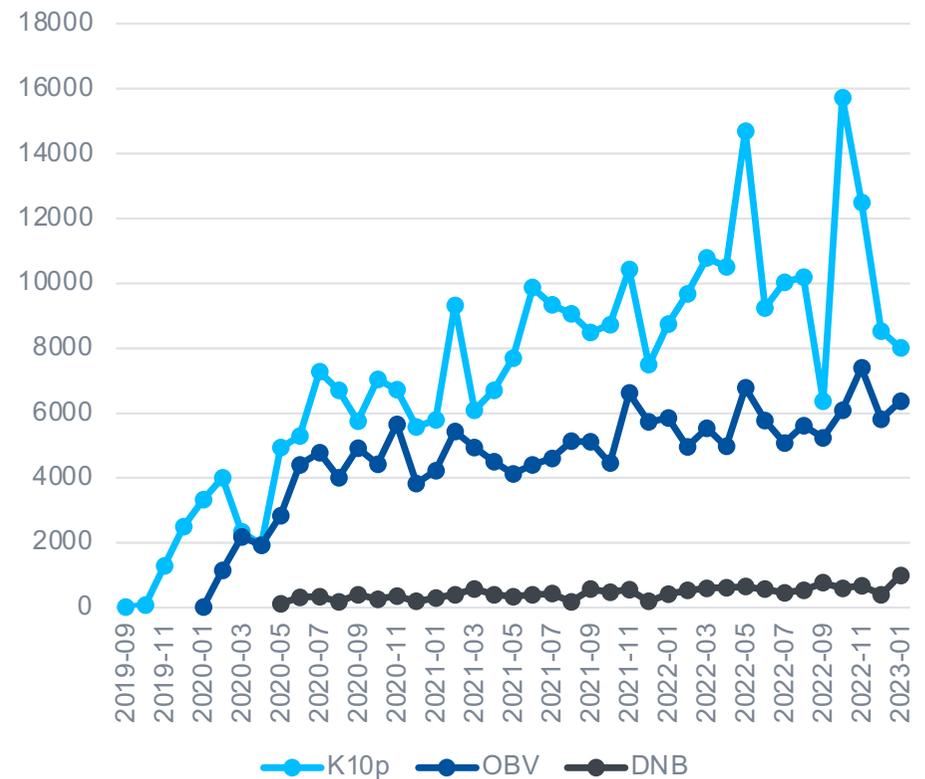


DA3-Nutzung 2019-2022 (bearbeitete Datensätze)

DA3-Nutzung Gesamt



DA3-Nutzungen pro Monat



Digitaler Assistent für Forschungsdaten (DA-FDM) - Projektüberblick



Projektpartner:

- UB Stuttgart (Projektleitung)
- BSZ
- Eurospider

Projektlaufzeit:

Start: 01.04.2023 (24 Monate)

Projektgenehmigung:

MWK Baden-Württemberg (02.02.2023)

Gefördert durch das MWK BW



Baden-Württemberg.de

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)



Universität Stuttgart

9. März 2023

Digitaler Assistent für Forschungsdatenmanagement

Der federführend von der Universität Stuttgart entwickelte und sehr erfolgreiche Digitale Assistent (DA) für Bücher, Zeitschriften und andere Medien wird weiterentwickelt. Das Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) fördert das neue Projekt DA-FDM, das zum Ziel hat, den Digitalen Assistenten auch für Forschungsdaten einzusetzen.

Der bereits entwickelte Digitale Assistent für Medien sorgt für eine effizientere Sacherschließung in Bibliothekssystemen. Der halbautomatische Assistent vereinfacht den Prozess, indem er Schlagworte vorschlägt und beschleunigt damit den Vorgang um das drei- bis vierfache. Er wurde seit 2015 federführend von der Universitätsbibliothek gemeinsam mit der schweizerischen Firma Europeizer entwickelt und ist sehr erfolgreich, wie sich unter anderem an der vielfachen Verwendung zeigt. Der Direktor der Universitätsbibliothek, Dr. Helge Steenweg, erklärt: „Der Digitale Assistent wird ausgehend von Baden-Württemberg mittlerweile in fast Deutschland, ganz Österreich, der Deutschen Nationalbibliothek (DNB) und der Staatsbibliothek zu Berlin eingesetzt.“

Schnelle und effiziente Erschließung von Forschungsdaten

Ziel des neuen Projekts DA-FDM ist nun, dass auch Forschungsdaten schneller und effizienter erschlossen werden können und dadurch weitaus nutzbarer sind. Für den DA-FDM sollen die Erkenntnisse aus den vorangegangenen Projekten genutzt werden. Als erster Schritt werden Forschungsdaten aus dem Fachbereich Ingenieurwissenschaften eingesetzt. Das MWK fördert das gemeinsam mit dem BSZ Konstanz | Bibliotheks-Service-Zentrum beantragte Projekt mit 500 000 Euro. Projektverantwortlicher ist Dr. Helge Steenweg.

Vereinfachte Handhabung beim Anlegen der Metadaten

<https://www.uni-stuttgart.de/universitaet/aktuelles/meldungen/Digitaler-Assistent-fuer-Forschungsdatenmanagement/>

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

Da in den Universitätsbibliotheken und Verbänden jahrzehntelange Erfahrung in der Sacherschließung von Monografien und Artikeln vorhanden ist und der Digitale Assistent (DA-3), der in weiten Teilen der BRD, in ganz Österreich sowie an der DNB im Einsatz ist, diese Erschließung erfolgreich seit Jahren durch halbautomatisierte Vorschläge unterstützt, wird in einem zweijährigen Projekt zwischen BSZ Konstanz, UB Stuttgart und der Firma Eurospider am Beispiel der Ingenieurwissenschaften der Prototyp eines Digitalen Assistenten für Forschungsdaten (DA-FDM) entwickelt werden.

Die standardisierte Erschließung wird die Metadatenerschließung sowohl für das Fachpersonal in den Bibliotheken und Forschungsdaten-Teams vereinfachen als auch Forschenden eine Hilfestellung bei der Abgabe der Forschungsdaten in Forschungsdatenrepositorien bieten.

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

Forschungsdaten können:

- als Addendum an Publikationen oder
- auch singulär als reine Forschungsdatensätze

vorkommen und werden entsprechend im Verbundkatalog und/oder in einem Repositoryum verzeichnet.

Grundidee:

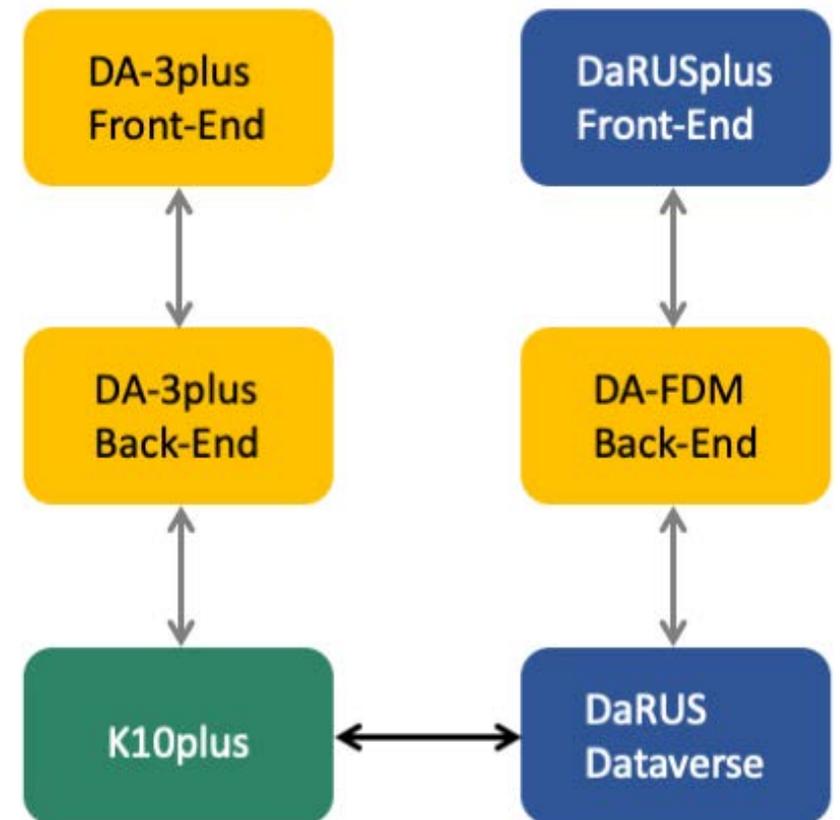
Erweiterung des DA um eine für Forschungsdaten relevante Komponente zur Abgabe von inhaltserschließenden Schlagwörtern für

- Forschende als auch
- bibliothekarisches Personal, sofern Forschungsdaten in den Bibliothekskatalogen z.B. durch Anhänge an Publikationen verzeichnet werden müssen

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

Da Forschungsdaten von der Genese unterschiedlich anfallen (allein oder als Anhang an Publikationen) soll der DA-FDM zweigleisig entwickelt werden :

In einem Teilstrang wird die geplante Funktionalität in das Front-End der bestehenden Repositorien-Software Dataverse an der UB Stuttgart integriert werden, um den Abgabeprozess für den Forschenden möglichst stringent ablaufen zu lassen. Im anderen Teilstrang wird der bestehende DA-3 ergänzt werden zu einem „DA-3plus“, mit dem dann in Bibliotheken Publikationen mit anhängenden Forschungsdaten erschlossen und in die Verbundkataloge (hier zunächst K10plus) angereichert werden können. Es wird zudem einen Metadaten-Abgleich zwischen Repository und Verbundkatalog geben.



Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

- **Anforderungen**
- das Forschungsdatenmanagement an Hochschulen wird immer wichtiger,
- lässt sich jedoch bezüglich der Metadatenbeschreibung vom einzelnen Bearbeitenden oder Forschenden nur relativ schwer umsetzen:
 - da die inhaltliche Erschließung sehr fachspezifisch durch die jeweils geltenden oder – wie in den Ingenieurwissenschaften – mit zu wenigen bzw. ungenügenden Ontologien zu realisieren ist.
- Um eine bessere Standardisierung zu erreichen, soll der DA-FDM in bewährter Technologie des DA-3 entsprechende Vorschläge machen, die dann halbautomatisch in den Verbundkatalog oder das Forschungsdatenrepositorium übernommen werden können.

(DA-FDM) – Forschungsdaten in Repositorien

- Je strukturierter und standardisierter Metadaten sind, umso einfacher können sie von Menschen und Maschinen zur Auffindbarkeit und interdisziplinären Verknüpfung genutzt werden.
- Im Rahmen der Nationalen Forschungsdateninfrastruktur (NFDI) entstehen zurzeit Terminologie-Services, Ontologien und Knowledge-Graph-basierte Infrastrukturen, die semantisch annotierte Terminologien für standardisierte Metadaten bereitstellen.
- Diese Technologien werden aktuell aber von Forschungsdatenrepositorien kaum zur inhaltlichen Erschließung der Daten und Codes genutzt.
- Die Beschreibung der Daten durch Metadaten obliegt in der Regel den Forschenden selbst.
- Eine manuelle Erschließung der Daten durch informationswissenschaftliches Fachpersonal ist aufwändig und findet daher nur in Ausnahmen statt.

DA-FDM – Forschungsdaten in Repositorien

<https://darus.uni-stuttgart.de/dataverse/darus>

The screenshot displays the DA-FDM (DataVerse) repository interface. At the top, the University of Stuttgart logo and name are visible, along with navigation links for Search, About, User Guide, Support, and Log In. Below the navigation bar, there is a search bar and a section for featured datasets. The featured datasets include SFB 1313 "Interface-Driven Multi-Field Processes in Porous Media", SFB 1333 "Molecular heterogeneous catalysis in confined geometries", and the Institute of Applied Mechanics (IAM). The main content area shows search results for "OrocoFEM data repository" and "Replication Data of Buchmeiser group for: Cationic Molybdenum Imido Alkylidene N-Heterocyclic Carbene Complexes Confined in Mesoporous Silica". The left sidebar contains filters for Dataset Category, Metadata Source, and Publication Year.

DA-FDM – Forschungsdaten in Repositorien

<https://darus.uni-stuttgart.de/dataverse/darus>

University of Stuttgart
Germany

SFB 1333 B6 - Peters group, IOC
(Universität Stuttgart)

DaFUS > SFB 1333 "Molecular heterogeneous catalysis in confined geometries" > SFB 1333 B6 - Peters group, IOC >

Replication Data for: Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxyde Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides

Version 1.1

Peters, Rene, 2023, "Replication Data for: Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxyde Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides", <https://doi.org/10.18419/darus-3014>, DaFUS, V1

[Cite Dataset](#) > [Learn about Data Citation Standards](#)

Description

development and optimization of catalysis, mechanistic experiments, synthetic description of ligand and catalyst synthesis, application of catalysts, synthetic use of products, spectroscopic characterization data of all new products, EPR data of complexes, X-ray data, biological activity of catalysis products and derivatives thereof, for interpretation see manuscript (2023-08-30)

Subject

Chemistry

Keyword

Catalyst Development (TEMA) <https://purl.org/tema/113291>
 Asymmetric Catalysis (TEMA) <https://purl.org/tema/113538>
 Reaction Mechanism (TEMA) <https://purl.org/tema/517811>
 Organic Synthesis (TEMA) <https://purl.org/tema/131654>
 EPR (Wikidata) <https://www.wikidata.org/wiki/Q260463>
 X-ray Crystal Structure Analysis
 Cell Painting Assay

Related Publication

Wanner, D. M., Becker, R. M., Suhr, S., Wannenmacher, N., Ziegler, S., Herrmann, J., Willig, F., Gabler, J., Jangid, K., Schmid, J., Hans, A. C., Frey, W., Sarkar, B., Klatner, J., Peters, R., Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxyde Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides, *Angew. Chem. Int. Ed.* 2023, 62, e202307317. doi: 10.1002/anie.202307317 <https://doi.org/10.1002/anie.202307317>

Funding Information

DFG: 356263783 - SFB 1333
 DFG: 210990693
 DFG: INST 40/575-1 FUGG (JUSTUS 2 cluster)
 Ministry of Science, Research and Arts Baden-Württemberg: bwHPC
 European Union: Drug Discovery Hub Dortmund (DDHD), EFRE-0200481
 European Union: Innovative Medicines Initiative (grant agreement number 115488)

Project

Cooperative Asymmetric Dual / Multiple Activation Catalysis Under Confinement (Level 0)

Depositor

Peters, Rene

Deposit Date

2023-07-24

License/Data Use Agreement

CC BY 4.0



(DA-FDM) – Forschungsdaten im Verbundkatalog

Es ist bereits heute möglich, Forschungsdaten in K10plus zu erfassen.

Durch die Erweiterung des DA3 durch Forschungsdaten hin zu einem „**DA3plus**“ sollen die vorhandenen Möglichkeiten zur Erfassung in K10plus erweitert und professionalisiert werden.

Der Nachweis der Metadaten zu Forschungsdaten im K10plus soll im Rahmen des Projekts deutlich ausgebaut werden.

Die Möglichkeit der automatisierten Übernahme der Metadaten aus FDM-Repositoryn soll zunächst mit dem Repository der Universität Stuttgart (DaRUS) über eine zu realisierende Schnittstelle geschaffen werden.

(DA-FDM) – Forschungsdaten im Verbundkatalog

The screenshot shows the BSZ GBV search interface. The search bar contains the text 'suchen [und]' and 'Suche über alles [ALL]'. The search results are sorted by 'Relevanz'. The main content area displays the following information for the search result:

K10plusPPN: 1858221188 **neu**

Titel: **MaiCuBeDa Hilprecht - Mainz Cuneiform Benchmark Dataset for the Hilprecht Collection** / Hubert Maria, Timo Homburg

Autorin/Autor: Maria, Hubert, 1975- [Verfasserin/Verfasser]

Beteiligt: Homburg, Timo [Verfasserin/Verfasser]

Erschienen: Heidelberg : Universität, 2023-08-30

Umfang: 1 Online-Ressource (11 Files)

Sprache(n): Englisch

Anmerkung: Gesehen am 30.08.2023

Bibliogr. Zusammenhang: Forschungsdaten zu: CNN based cuneiform sign detection learned from annotated 3D renderings and mapped photographs with illumination augmentation / Stötzner, Ernst [GNDNR:1300462698], 22 Aug 2023
 Forschungsdaten zu: R-CNN based polygonal wedge detection : learned from annotated 3D renderings and mapped photographs of ppen data Cuneiform tablets / Stötzner, Ernst [GNDNR:1300462698], 2023
 Forschungsdaten zu: 3D data derivatives of the Haft Tappelh processing pipeline / Homburg, Timo [GNDNR:1300461551], 2022-07-20

Grant number: 424957759, DFG

Link zum Volltext: Elektronische Ressource: Zugang beim Produzenten [Lizenzangabe: **Lizenzpflichtig**]
 Elektronische Ressource: Zugang beim Produzenten [Lizenzangabe: **Lizenzpflichtig**]
 Digital Object Identifier [DOI]: 10.11588/data/Q5N1Q2

Art und Inhalt: Forschungsdaten | Datenbank | Computerdaten

Sonstige Schlagwörter: Arts and Humanities ; Computer and Information Science

Inhaltliche Zusammenfassung: Das Mainz Cuneiform Benchmark Dataset (MaiCuBeDa) beinhaltet Bilder von Keilschriftzeichen, Worten bestehend aus Keilschriftzeichen, Keilschriftzeichenzellen und annotierten Einzelzeilen basierend auf dem Datenset MaiCuBeDa Hilprecht: <https://doi.org/10.11588/data/TE8CCN>. Die Annotationen werden in drei verschiedenen Renderings mit Metadaten in CSV und einem Knowledge Graph (RDF) bereitgestellt.

Besitzende Bibliotheken: [Heidelberg, Hei8B - Die Heidelberger Universitätsbibliothek <16/250>](#)
[Heidelberg, Universitätsbibliothek Heidelberg <16>](#)

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

Im Rahmen dieses Projektantrages wird der Fokus zunächst auf die ingenieurwissenschaftliche Fachdisziplin begrenzt, um beispielhaft die Funktionalität und die Anwendungsmöglichkeiten einer halbautomatischen Metadaten-Erschließung zu überprüfen.

Für FD-Repositoryn wird zunächst eine Entwicklung des DA-FDM-Back-Ends für die Repositoryn-Software DataVerse angestrebt.

Digitaler Assistent für das Forschungsdatenmanagement (DA-FDM)

Das Fachpersonal in Bibliotheken hat die Möglichkeit, mit einem erweiterten Digitalen Assistenten (DA-3plus Front-End) Korrekturen und Erweiterungen anzubringen bzw. Publikationen mit anhängenden Forschungsdaten zu erfassen.

Vorgesehen ist auch hier eine Auswahl von automatisch vorgeschlagenen Notationen aus geeigneten Klassifikationen.

Eine Ähnlichkeits-Suchfunktion soll zusätzlich Vorschläge bieten.

Die so erschlossenen Titel stehen über ein DA-3plus Back-End dann allen K10plus-Teilnehmern zur Verfügung und können über verschiedene Schnittstellen ausgeliefert werden.

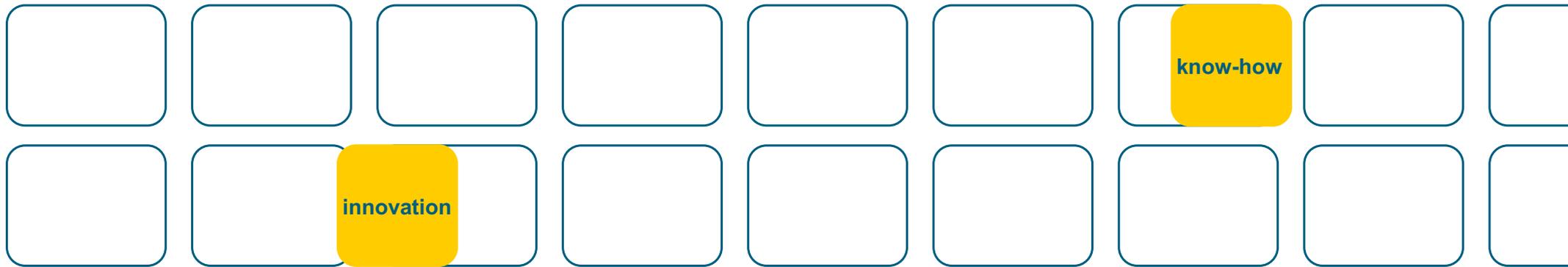
DA-FDM Projektteam

UB Stuttgart	BSZ	Eurospider
Karoline Weinspach (UB, Projekt)	Volker Conradt	Dr. Peter Schäuble
Dr. Helge Steenweg (UB)	Kim Hornung	Dr. Gérard Milmeister
Dr. Imma Hinrichs (UB)	Silke Horny	Thomas Murphy
Dr. Dorothea Iglezakis (UB-Fokus)	Armin Kühn	
Ph.D. Sarbani Roy (UB-Fokus, Projekt)	N.N. (Projekt)	

DA-FDM – Mitarbeit, Community

Um auch Feedback von den späteren Nutzern zu erhalten, sollen zwei Workshops mit Forschenden, FD-Repository-Betreibenden und bibliothekarischen Fachleuten je nach Fertigstellung der beiden geplanten Prototypen Aufschlüsse und Anregungen zur Nutzbarkeit bringen.

Mitarbeit ist bei Interesse sehr erwünscht



DA-FDM – Ein Digitaler Assistent zur halbautomatischen Erschliessung von Forschungsdaten

- 2. November 2023, Helge Steenweg & Peter Schäuble

Uebersicht

- Erste Erkenntnisse nach Projektstart
- Worum geht es bei der Inhaltsermittlung von Forschungsdaten?
- Was sind die technischen Herausforderungen
- Wie wollen wir diese meistern
- Zusammenfassung

Was sind Forschungsdaten?

Was ist falsch?

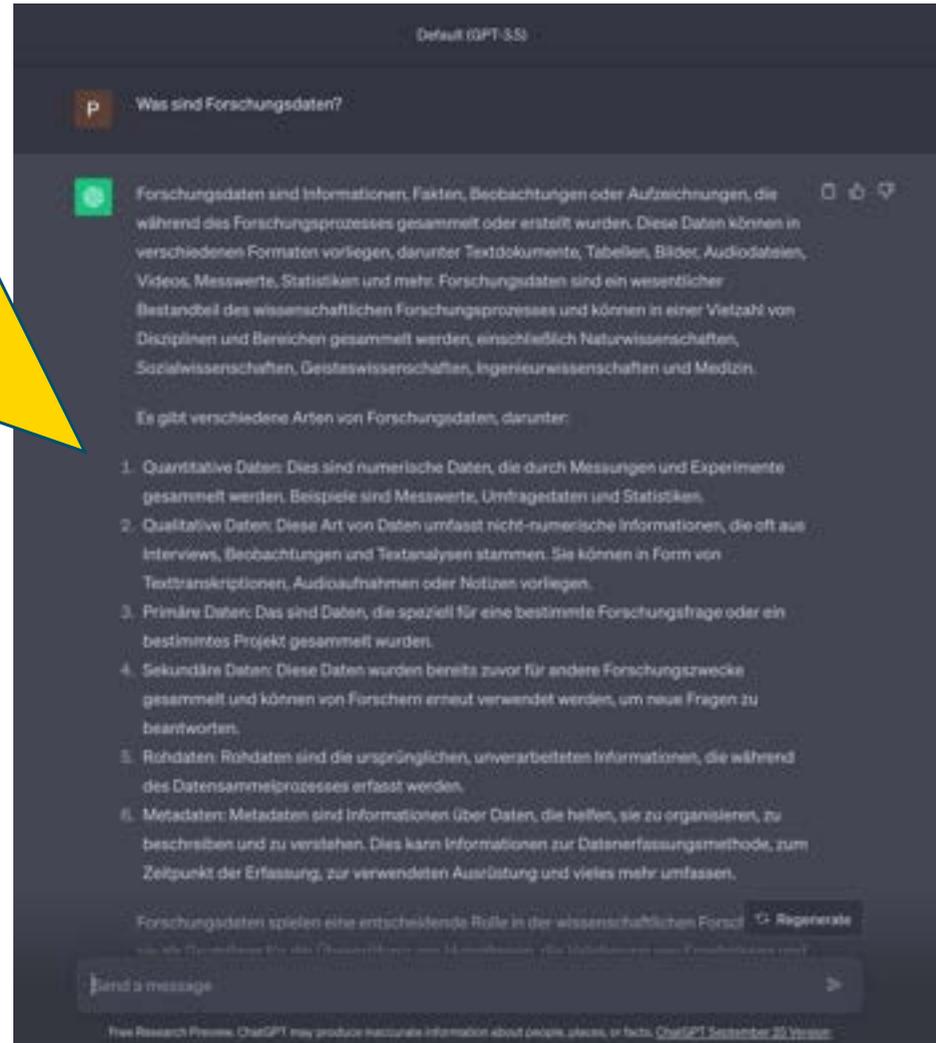
Primärdaten = Ausgangsdaten;
Sekundärdaten = Ergebnisdaten
(<https://de.wikipedia.org/wiki/Forschungsdaten>)

Was fehlt?

Wissenschaftliche Publikationen,
welche mit den und/oder über
die FD verfasst wurden und
Applikationsdaten (ganze
Datenbankexporte)

Replik, kritische Würdigung?

Überschneidung mit github?
Kein Publikationsprozess mit Peer
Review etc.



The screenshot shows a chat interface with a dark theme. At the top, it says "Default (GPT-3.5)". The user's question is "Was sind Forschungsdaten?". The AI's response is in German and defines research data as information, facts, observations, or recordings collected during the research process. It lists various formats like text documents, tables, images, audio files, videos, and statistics. It also lists six types of research data: 1. Quantitative data (numerical, from measurements/experiments), 2. Qualitative data (non-numerical, from interviews/observations), 3. Primary data (collected for a specific purpose), 4. Secondary data (previously collected for other purposes), 5. Raw data (original, unprocessed information), and 6. Metadata (information about data). At the bottom, there is a "Send a message" input field and a "Regenerate" button. A small disclaimer at the very bottom reads: "Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. ChatGPT, September 26, 2023."

Braucht es Inhaltserchließung?

Ziel der inhaltlichen Erschließung ist, das Auffinden relevanter Ressourcen [...] zu erleichtern und zu beschleunigen. Der Informationswissenschaftler nennt dies „das **Retrieval verbessern**“ (<https://de.wikipedia.org/wiki/Sacherschlie%C3%9Fung>)

- Aus dem Blickwinkel des Ordnungsprinzips **Gleiches-zu-Gleichem** definieren Gödert, Lepsky und Nagelschmidt (2011, S. 70) als Ziel der Inhaltserchließung, dass „[...] ein Dokument bei inhaltlich gleichen – oder zumindest ähnlichen – Dokumenten eingeordnet werden kann“
- Fugmann definiert die inhaltliche Erschließung als eine das Wiederfinden erleichternde Maßnahme, „[...] ohne daß hierbei untragbar große Kosten entstehen.“ (1999, S. 11). Der Einsatz einer Erschließungsmethode ist somit ein Kompromiss, der den Nutzen und die Kosten der eingesetzten Methode im Kalkül berücksichtigt.
- Eine Studie der IDC zeigt auf, dass ein typisches Unternehmen mit 1000 Wissensarbeitern (engl. knowledge workers), die 30% ihrer Arbeitszeit (2,5 h/Tag) mit der Suche nach Informationen verbringen (Vgl. Feldman & Sherman, 2001, S. 7f.)

Zuviel vs. zuwenig erschliessen?

Zuwenig erschliessen

Die USA waren 1957 ein halbes Jahr lang mit dem Dechiffrieren des **Sputnik-Signals** beschäftigt, obwohl die verwendeten Signale und ihre Verschlüsselung lange vor dem Start des Satelliten von amerikanischen Stellen übersetzt und **in heimischen Bibliotheken** verfügbar waren (vgl. Rauch, 1988, S. 8f., zit. nach: Stock, 2006, S. 40)

Wer soll wieviel erschliessen?



Bibliothek von Alexandria

Rudolf Blum konstruiert aus dieser und weiteren fragmentarischen Überlieferungen schließlich die folgende Gesamtsystematik, die einen Einblick in die Akkuratessse der Literaturverzeichnung der Alexandriner gibt.

- Kallimachos teilte zunächst die griechischen Autoren nach Klassen ein (Rhetorik, Recht, Epik, Tragik, Komik, Lyrik, Geschichte, Medizin, Mathematik, Naturwissenschaft, Vermischtes) und bildete innerhalb dieser ggf. Unterabteilungen;
- innerhalb der Klassen bzw. Abteilungen nahm er eine alphabetische Sortierung der Autoren vor;
- wo es ihm möglich war, fügte er den Namen der Autoren biographische Daten hinzu;
- unter der Autorenangabe verzeichnete er schließlich sämtliche Schriften desselben und ordnete sie wiederum in Kategorien;
- gab er kurz die ersten paar Worte jedes aufgeführten Werkes (Incipit) wieder und
- den Umfang (die Anzahl στίχοι).



Welches Erschliessungs- bzw. Ordnungssystem für Forschungsdaten?

Provisorische Übersicht: Erschliessungssysteme für FDM in Deutschland
(Roy, 9. Okt. 2023).

- RWTH Aachen: DDC Subject classification
- RADAR (KIT): Standardized GND entry or free text. (GND or other)
- DepositOnce (TU Berlin): DDC Subject classification. (DDC or other). No con-vocab for keywords.
- TU Braunschweig: DDC Subject classification. No con-vocab for keywords.
- TU Darmstadt: DFG and DDC subject classification. No con-vocab for keywords.
- TU Dresden: No con-vocab for keywords.
- TU Munich: TUM Classification System. No con-vocab for keywords.
- LMU Hannover: No con-vocab for keywords. No classification found.
- Fraunhofer: DDC classification

Anforderungen an Erschliessungssystem für Forschungsdaten

- Muss auch neue Forschungsgebiete abdecken
- Muss mit wachsender Menge von Forschungsdaten skalieren
- Mehrsprachig (min. EN und DE)
- Trainingsdaten für (halb-)automatische Erschliessung
- Katalogstrukturen vorhanden
- Know How bei Forschenden und Fachreferenten vorhanden
- langfristige Pflege sichergestellt (Liste der nicht gepflegten Systeme ist gross)
- frei verfügbar, keine Lizenzabgaben

ERSCHLIESSUNGSSYSTEME I

* GND

Anzahl Sachschlagwörter:
~204'000
Sprache: Deutsch
Verweise: LCSH, DDC

* RVK

Anzahl Schlagwörter: 674'314
Sprache: Deutsch
Verweise: GND

* BK

Anzahl Klassen: 2087
Sprache: Deutsch

* IxTheo

Anzahl Klassen: 139
Domäne: Theologie
Sprache: Deutsch, Englisch

* MeSH

Anzahl Deskriptoren: 30'194
Domäne: Medizin
Sprache: Englisch

* NLM

Anzahl Codes: 4'702
Domäne: Medizin
Sprache: English

* LCC

Anzahl Klassen: ~650'000
Sprache: English

ERSCHLIESSUNGSSYSTEME II

* EuroThes

Anzahl Schlagwörter: 10'702
Sprache: Deutsch, English,
Französisch, ...

* STW

Anzahl Schlagwörter: 7'570
Domäne: Wirtschaft
Sprache: Deutsch, Englisch
Verweise: GND

* DDC Sachgruppen

Anzahl Klassen: 110
Sprache: Deutsch
Verweise: DDC

* FAST Topical

Anzahl Schlagwörter: ~460'000
Sprache: Englisch
Verweise: LCSH

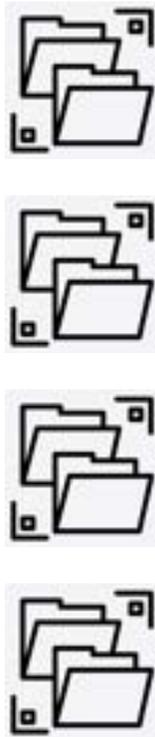
* MSC

Anzahl Schlagwörter: 6'198
Domäne: Mathematik
Sprachen: English (6198),
Italienisch (4755), Russisch (186),
Chinesisch (6195)

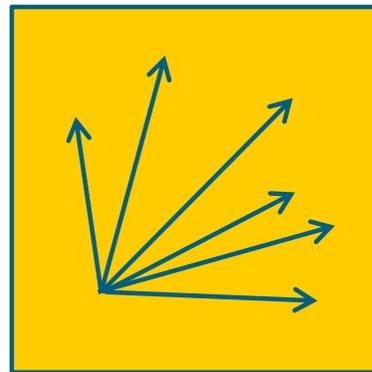
* TEMA

Anzahl Schlagwörter: 59'456
Domäne: Technik und
Management
Sprache: Englisch (mit deutschen
"Synonymen")

Forschungsdatensätze



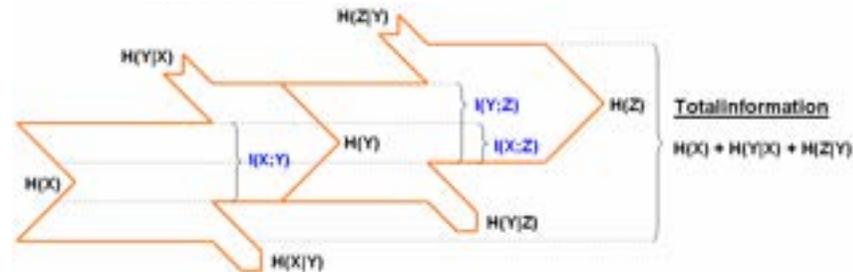
KI generierte Vektor-Repräsentanten



bibliothekarische Ordnungssystem



Zusatzinformation (Fehlinformation des Kanals)



Äquivokation (Verlustentropie)

Zusammenfassung

- Mengengerüst
 - Bibliothek von Alexandria: 500'000
 - DaRUS UB Stuttgart: 2244 Datasets mit 127.858 Dateien
 - e-collection ETH Zürich: 2000
 - jährliches Wachstum: 50 bis 100%
- Was sind die Herausforderungen?
 - unterschiedliche Anforderungen der verschiedenen Akteure (Innen-/Aussensicht)
 - geeignetes Erschliessungssystem & Compatibilität mit anderen Repositorien
 - Künstliche Intelligenz vs. etablierte Prozesse
- Nächste Schritte
 - KI generiertes Ordnungssystem
 - Vorschlagsgenerator aufgrund von Publikationen und inhaltlicher Zusammenfassung

FACHVERANSTALTUNG

7. WORKSHOP COMPUTERUNTERSTÜTZTE INHALTSERSCHLIESSUNG 2023

Beginn 14.11.2023, 14:00 Uhr

Ende 15.11.2023, 13:30 Uhr

Wo Virtuell

[↓ In meinen Kalender eintragen \(ICS, 2kB\)](#)



Vielen Dank!



Dr. Helge Steenweg

Holzgartenstr. 16
70174 Stuttgart

ub.uni-stuttgart.de/ubs/personen/Steenweg-00002/



Dr. Peter Schäuble

Winterthurerstrasse 92
8006 Zürich

<https://www.eurospider.com/de/ueber-uns>