

Universität Stuttgart

BSZ Bibliotheksservice-Zentrum
Baden-Württemberg

 **eurospider**
relevancy retrieval

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

Dr. Peter Schäuble, Dr. Helge Steenweg

15. 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

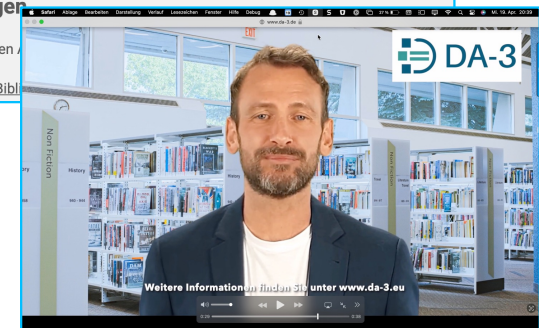


Baden-Württemberg.de

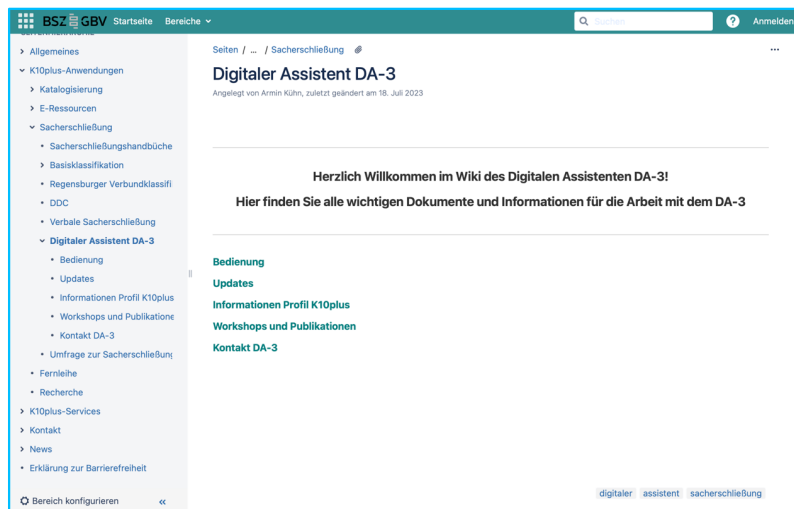
The screenshot shows the DA-3 website homepage. At the top right, there are links for 'Wiki: DA-3-Wiki' and 'Support: E-Mail schreiben'. The main header features the DA-3 logo and navigation links: 'Über den DA-3', 'Anwender*in werden', and 'FAQ'. The main content area has a heading: 'Der Digitale Assistent DA-3 ist eine Plattform zur Unterstützung der inhaltlichen Erschließung in Bibliotheken'. Below this, it lists '3 gute Gründe für die Arbeit mit dem DA-3':

- One-Stop-Solution für die Inhaltserschließung nutzen
- Quantität und Qualität der kooperativen Inhaltserschließung steigern
- Thematische Recherche in Discovery-Systemen und Katalogen optimieren

A paragraph follows: 'Viele Einrichtungen in Deutschland und Österreich verwenden den DA-3 bereits, um die intellektuelle Inhaltserschließung effizient zu gestalten, Synergieeffekte zu nutzen und den Aufwand zu reduzieren. Auf dieser Webseite finden Sie die wichtigsten Informationen rund um die Arbeit mit dem DA-3 und für Interessierte erste Schritte zur Teilnahme.' At the bottom, there are sections for 'NEWS' (with a sub-heading 'Beim Update am 14.02.2023 wurden neue Listenfunktionen aufgeschaltet.') and 'Veranstaltungen' (with a sub-heading 'Vorträge zum Digitalen Rahmen des 1. Österreichischen Bibliographischen Kongresses').



Verbreitung des Digitalen Assistenten für die halbautomatische Sacherschließung



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

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

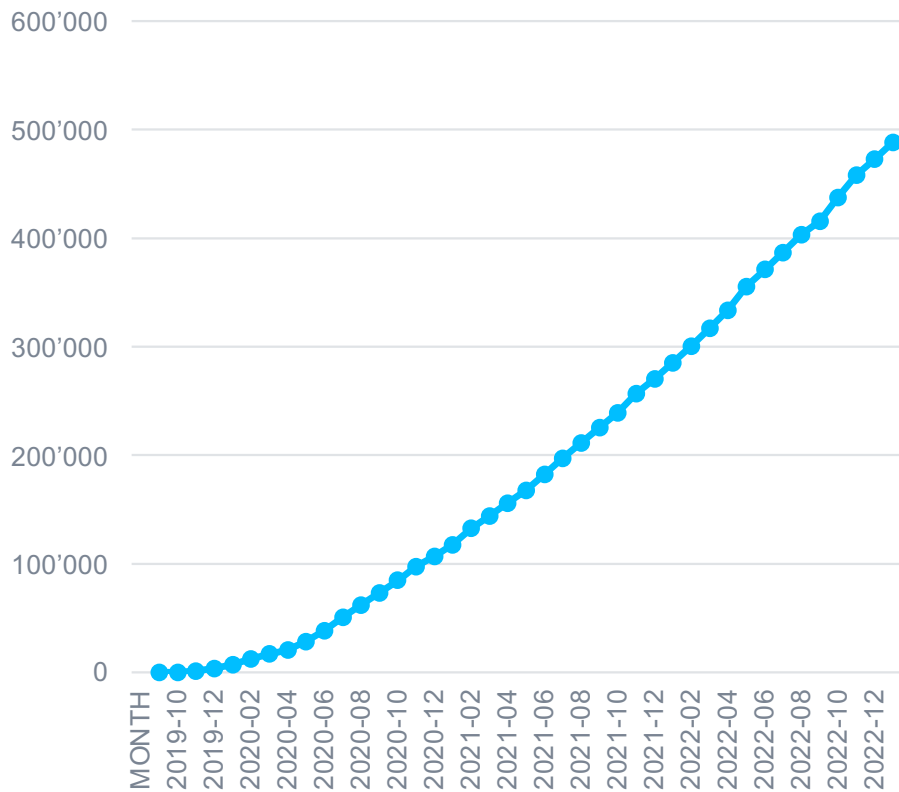
<https://www.bsz-bw.de/kooperationen/da-3.html>

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

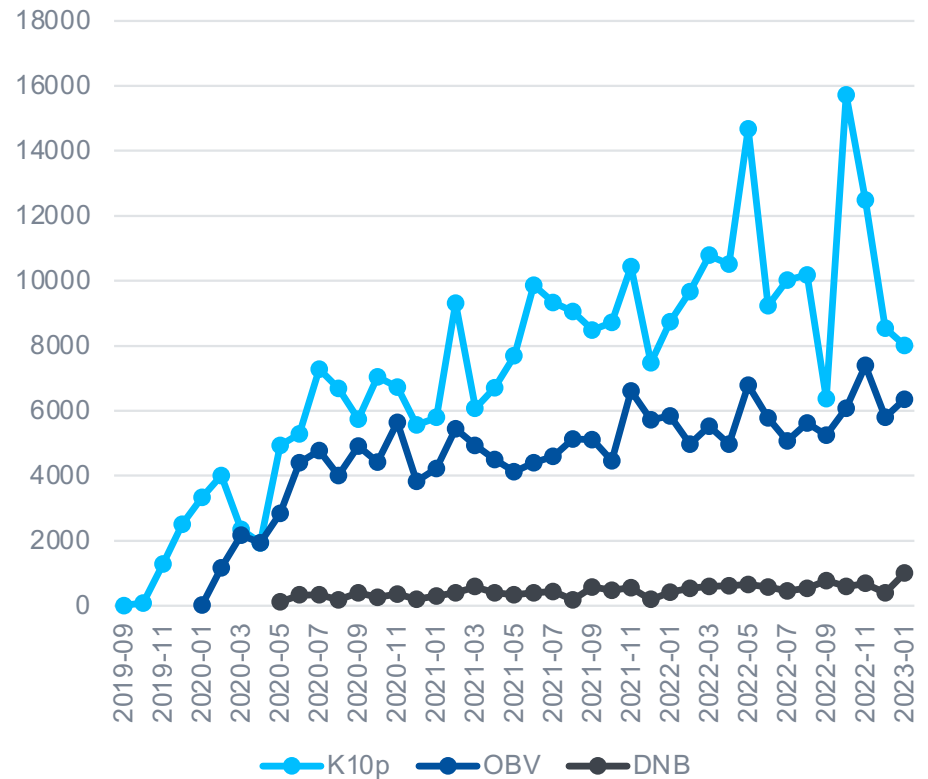


DA3-Nutzung 2019-2022 (bearbeitete Datensätze)

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)

Forschungsdaten können:

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

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

Grundidee:

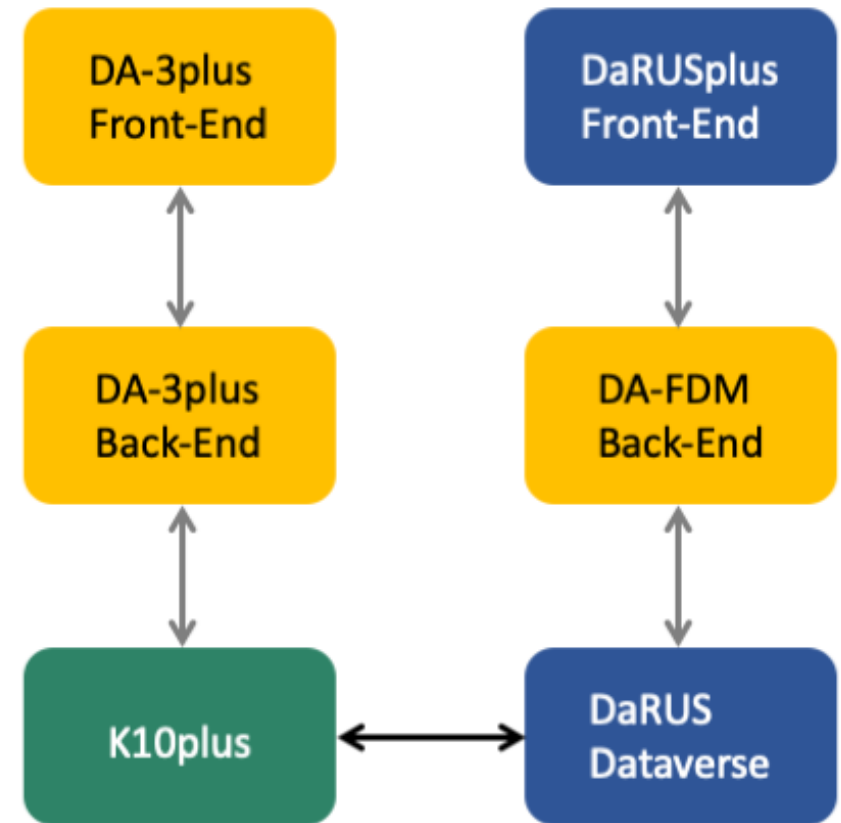
Erweiterung des DA um eine für Forschungsdaten relevante Komponente zur Abgabe von inhaltserschließenden Schlagwörtern für

- 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.



(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 website interface. At the top, the University of Stuttgart logo and name are visible, along with the DaRUS logo. A navigation bar includes links for Search, About, User Guide, Support, and Log In. Below the navigation bar, a metrics section shows 777,053 Downloads. The main content area features a carousel of featured research groups: SFB 1313, SFB 1333, Institute of Biochemistry and Technical Biochemistry, and Institute of Applied Mechanics (MIB). A search bar is present with a search icon and a link to Advanced Search. The search results section shows 1 to 10 of 2,031 results. On the left, there are filters for Dataverses (669), Datasets (1,362), and Files (70,463). Below these are filters for Dataverse Category (Research Project, Research Group, Organization or Institution, Department, Laboratory) and Metadata Source (DaRUS, Harvested). The search results list three items, each with a document icon, title, date, and a brief description of the data repository or replication data.

University of Stuttgart
Germany

Search ▾ About User Guide Support Log In

Metrics 777,053 Downloads Contact Share

This is the data Repository of the University of Stuttgart.

SFB 1313 "Interface-Driven Multi-Field Processes in Porous Media"

SFB 1333 "Molecular heterogeneous catalysis in confined geometries"

Institute of Biochemistry and Technical Biochemistry

Institute of Applied Mechanics (MIB)

Search this dataverse... Advanced Search

Dataverses (669)
Datasets (1,362)
Files (70,463)

Dataverse Category
Research Project (480)
Research Group (75)
Organization or Institution (64)
Department (10)
Laboratory (9)
More...

Metadata Source
DaRUS (2,026)
Harvested (5)

Publication Year
2023 (222)
2022 (277)
2021 (1,372)
2020 (107)
2019 (48)

1 to 10 of 2,031 Results Sort ▾

OncoFEM data repository
Sep 19, 2023 - Institute of Structural Mechanics and Dynamics in Aerospace Engineering
Suditsch, Marlon; Ricken, Tim; Wagner, Arndt, 2023, "OncoFEM data repository", <https://doi.org/10.18419/darus-3679>, DaRUS, V1
In this repository the data files of OncoFEM are collected. These are in the following listed and cited where appropriate: SRI24 atlas T1 and T2 MRI modalities (T. Rohlfing, N. M. Zahr, E. V. Sullivan, and A. Pfefferbaum, "The SRI24 multichannel atlas of normal adult human brain...

Replication Data of Buchmeiser group for: Cationic Molybdenum Imido Alkylidene N-Heterocyclic Carbene Complexes Confined in Mesoporous Silica: Tuning Transition States Towards Z-Selective Ring-Opening Cross-Metathesis
Sep 15, 2023 - SFB 1333 A1+B2 - Buchmeiser group, IPOC-MSF
Goldstein, Elizabeth L.; Buchmeiser, Michael; Ziegler, Felix; Beurer, Ann-Katrin; Traa, Yvonne; Bruckner, Johanna R., 2023, "Replication Data of Buchmeiser group for: Cationic Molybdenum Imido Alkylidene N-Heterocyclic Carbene Complexes Confined in Mesoporous Silica: Tuning Transition States Towards Z-Selective Ring-Opening Cross-Metathesis", <https://doi.org/10.18419/darus-3696>, DaRUS, V1
All primary data files of measurements and processed data of the journal article mentioned under related publications from the Buchmeiser group can be found here. The data is structured according to figures and schemes in the research article and contains the following data types...

Replication Data for: Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxide Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides
Sep 14, 2023 - SFB 1333 B6 - Peters group, IOC
Peters, Rene, 2023, "Replication Data for: Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxide Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides", <https://doi.org/10.18419/darus-3614>, DaRUS, V1
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...

DA-FDM – Forschungsdaten in Repositorien

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

University of Stuttgart
Germany

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

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

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

Version 1.0

Peters, Rene, 2023, "Replication Data for: Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxide Catalysis: Pyrazolone Addition to Nitroolefins as Entry to Diaminoamides", <https://doi.org/10.18419/darus-3614>, DaRUS, 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, Asymmetric Catalysis, Reaction Mechanism, Organic Synthesis, EPR, X-ray Crystal Structure Analysis, Cell Painting Assay

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

License/Data Use Agreement CC BY 4.0

University of Stuttgart
Germany

Search About User Guide Support Log In

Persistent Identifier doi:10.18419/darus-3614

Publication Date 2023-09-14

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

Author Peters, Rene (Universität Stuttgart) - ORCID: 0000-0002-6668-4017

Point of Contact Use email button above to contact.
Peters, Rene (Universität Stuttgart)

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/017811>
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, P. M., Suhr, S., Wannemacher, N., Ziegler, S., Herrmann, J., Willig, F., Gabler, J., Jangid, K., Schmid, J., Hans, A. C., Frey, W., Sarkar, B., Kästner, J., Peters, R., Cooperative Lewis Acid-1,2,3-Triazolium-Aryloxide 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: 358283783 - SFB 1333
DFG: 310990893
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 115489)

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

Depositor Peters, Rene

Deposit Date 2023-07-24

(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-Repositoryen soll zunächst mit dem Repositoryum der Universität Stuttgart (DaRUS) über eine zu realisierende Schnittstelle geschaffen werden.

(DA-FDM) – Forschungsdaten im Verbundkatalog

BSZ GBV

[Einfache Suche](#) |
 [Erweiterte Suche](#) |
 Suchergebnis |
 [Zwischenablage](#) |
 [Hilfe](#)

[Impressum](#) |
 [Datenschutz](#) |
 © 2023 OCLC

K10plus
Verbundkatalog
GBV|SWB

sortiert nach

nur Zeitschriften/Serien/Datenbanken
 nur Online-Ressourcen
 OpenAccess
 Unschärfe Suche

4 von 4647
4 von 4647

◀ ▶
Ihre Aktion: bezogen auf Schlagwort: Forschungsdaten
◀ ▶

Felder \
ISBD \
MARC21 (FL_924) \
Citavi, Referencemanager (RIS) \
Endnote Tagged Format \
BibTex-Format \
RDF-Format \

K10plusPPN:
 Titel:
 Autorin/Autor:
 Beteiligt:
 Erschienen:
 Umfang:
 Sprache(n):
 Anmerkung:
 Bibliogr. Zusammenhang:

1858221188 [Filter](#)

MaiCuBeDa Hilprecht - Mainz Cuneiform Benchmark Dataset for the Hilprecht Collection / Hubert Mara, Timo Homburg [Zum Volltext](#)

Mara, Hubert, 1975- [Verfasserin/Verfasser] [↗](#)

Homburg, Timo [Verfasserin/Verfasser] [↗](#)

Heidelberg : Universität, 2023-08-30

1 Online-Ressource (11 Files)

Englisch

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 Tappeh processing pipeline / Homburg, Timo [GNDNR:1300461551], 2022-07-20 424957759, DFG

Grant number:

Link zum Volltext: [Elektronische Ressource: Zugang beim Produzenten \(Lizenzangabe: **Lizenzpflichtig**\)](#)
 Elektronische Ressource: Zugang beim Produzenten (Lizenzangabe: **Lizenzpflichtig**)
 Digital Object Identifier (DOI): 10.11588/data/QSNIQ2

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, Keilschriftzeichenzeilen und annotierten Einzelkeilen basierend auf dem Datenset HeiCuBeDa Hilprecht: <https://doi.org/10.11588/data/IE8CCN> . Die Annotationen werden in drei verschiedenen Renderings mit Metadaten in CSV und einem Knowledge Graph (RDF) bereitgestellt.

4 von 4647
4 von 4647

◀ ▶

Besitzende Bibliotheken
Detailsansicht

Heidelberg, HeiBIB - Die Heidelberger Universitätsbibliographie <16/250>
 Heidelberg, Universitätsbibliothek Heidelberg <16>

Universitätsbibliothek Stuttgart | DA3-Workshop | Ste | 15. Nov. 2023 | Folie 12

12

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.

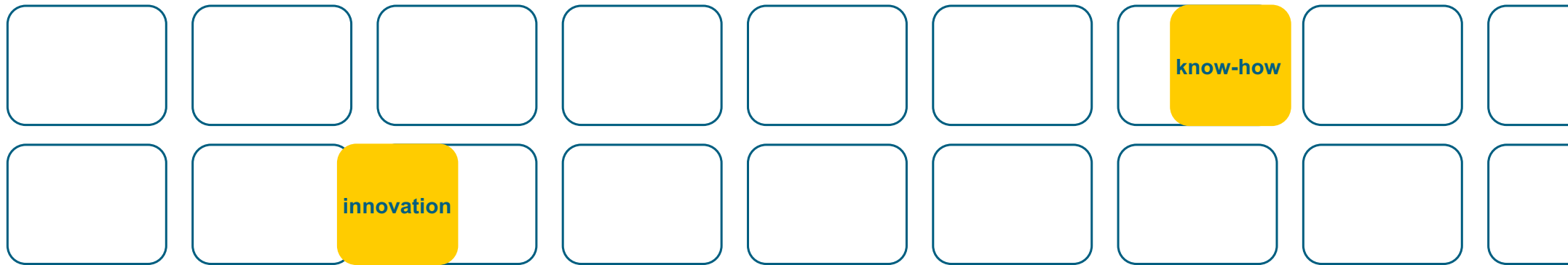
Vorgesehen ist auch hier eine Auswahl von automatisch vorgeschlagenen Notationen aus geeigneten Klassifikationen.

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 – 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.



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

15. November 2023, Helge Steenweg & Peter Schäuble

Übersicht

- Erste Erkenntnisse nach Projektstart
- Worum geht es bei der Inhaltserschließung 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.

Default (GPT-3.5)

P Was sind Forschungsdaten?

Forschungsdaten sind Informationen, Fakten, Beobachtungen oder Aufzeichnungen, die während des Forschungsprozesses gesammelt oder erstellt wurden. Diese Daten können in verschiedenen Formaten vorliegen, darunter Textdokumente, Tabellen, Bilder, Audiodateien, Videos, Messwerte, Statistiken und mehr. Forschungsdaten sind ein wesentlicher Bestandteil des wissenschaftlichen Forschungsprozesses und können in einer Vielzahl von Disziplinen und Bereichen gesammelt werden, einschließlich Naturwissenschaften, Sozialwissenschaften, Geisteswissenschaften, Ingenieurwissenschaften und Medizin.

Es gibt verschiedene Arten von Forschungsdaten, darunter:

1. Quantitative Daten: Dies sind numerische Daten, die durch Messungen und Experimente gesammelt werden. Beispiele sind Messwerte, Umfragedaten und Statistiken.
2. Qualitative Daten: Diese Art von Daten umfasst nicht-numerische Informationen, die oft aus Interviews, Beobachtungen und Textanalysen stammen. Sie können in Form von Texttranskriptionen, Audioaufnahmen oder Notizen vorliegen.
3. Primäre Daten: Das sind Daten, die speziell für eine bestimmte Forschungsfrage oder ein bestimmtes Projekt gesammelt wurden.
4. Sekundäre Daten: Diese Daten wurden bereits zuvor für andere Forschungszwecke gesammelt und können von Forschern erneut verwendet werden, um neue Fragen zu beantworten.
5. Rohdaten: Rohdaten sind die ursprünglichen, unverarbeiteten Informationen, die während des Datensammelprozesses erfasst werden.
6. Metadaten: Metadaten sind Informationen über Daten, die helfen, sie zu organisieren, zu beschreiben und zu verstehen. Dies kann Informationen zur Datenerfassungsmethode, zum Zeitpunkt der Erfassung, zur verwendeten Ausrüstung und vieles mehr umfassen.

Forschungsdaten spielen eine entscheidende Rolle in der wissenschaftlichen Forschung. Sie sind die Grundlage für die Überprüfung von Hypothesen, die Validierung von Ergebnissen und

Send a message

Regenerate

Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. ChatGPT September 25 Version

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 (z.B. wegen zeitgemässen Bezeichnungen)
- 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: Englisch

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

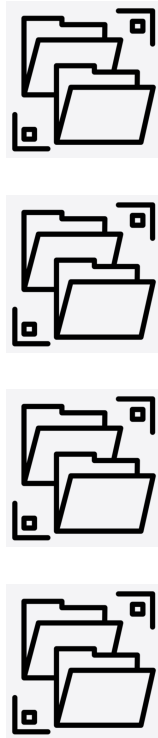
* MSC

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

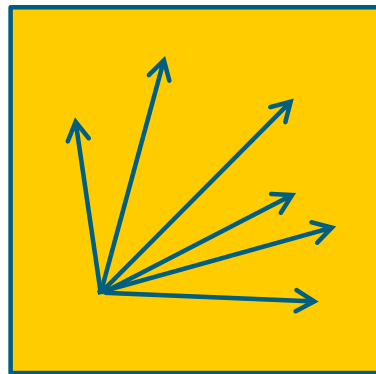
* TEMA

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

Forschungs-
datensätze



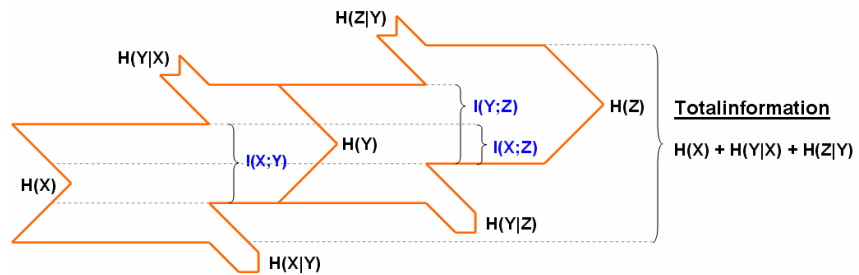
KI generierte
Vektor-Repräsentanten



bibliothekarische
Ordnungssystem



Zusatzinformation (Fehlinformation des Kanals)



Äquivokation (Verlustentropie)

| BK | | MESH | |
|-------|-------------------------|---|-------|
| 44.36 | Embryologie | Embryology | 5'627 |
| 44.35 | Histologie | Histology | 5'452 |
| 44.61 | Innere Medizin | Internal Medicine | 5'159 |
| 44.62 | Allgemeinmedizin | Family Practice | 3'891 |
| 44.34 | Anatomie | Anatomy | 1'832 |
| 44.37 | Physiologie | Physiology | 1'450 |
| 44 | Medizin | Medicine | 1'289 |
| 44.47 | Pathologie | Pathology | 1'282 |
| 35.7 | Biochemie: Allgemeines | Biochemistry | 1'192 |
| 44.51 | Diagnostik | Medical History Taking Physical Examination | 1'085 |
| 44.7 | Sportmedizin | Sports Medicine | 1'075 |
| 44.06 | Medizinsoziologie | Sociology, Medical | 1'038 |
| 44.69 | Intensivmedizin | Intensive Care | 1'022 |
| 44.72 | Rechtsmedizin | Forensic Medicine | 993 |
| 44.67 | Kinderheilkunde | Pediatrics | 895 |
| 44.35 | Histologie | Atlases Histology | 847 |
| 44.66 | Anästhesiologie | Anesthesia | 845 |
| 44.78 | Immunkrankheiten | Hypersensitivity | 820 |
| 44.68 | Gerontologie, Geriatrie | Geriatrics | 801 |
| 44.48 | Medizinische Genetik | Genetics, Medical | 799 |

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

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>