

Franziska Fuchs

Das DNBLab

Inhaltsverzeichnis

- **Was ist Text- und Data Mining?**
- **Exemplarische Beispiele**
- **Was ist das DNBLab?**
- **Was bietet das DNBLab?**
- **Ausflug ins DNBLab**
- **Ausblick**

Was ist Text und Data Mining?

- computergestützte Analyse großer Datenmengen zur Beantwortung wissenschaftlicher Fragestellungen
- Daten u.a. Texte (Monografien, Periodika, Inhaltsverzeichnisse), Bilder, Musik
- Nutzung in Wissenschaft, Industrie und Gesellschaft
- Zweck: Entwicklung neuer wissenschaftlicher Thesen & Überprüfung alter Thesen

Exemplarische Beispiele

- Wortentstehungsgeschichte
- Herausfinden aktueller Publikationstrends
- Demografische oder politische Untersuchungen
- Themen- und Stimmungsanalysen

Was ist das DNBLab?

- Bibliotheks-Lab der Deutschen Nationalbibliothek
- Zentraler Anlaufpunkt für Zugang und Analyse der freien digitale Objekte der DNB
- Start im Sommer 2020

DNBLAB NEWSLETTER

Mit unserem Newsletter informieren wir Sie über Veranstaltungen und Neuigkeiten aus dem DNBLab.

Mehr 

Was bietet das DNBLab?

- freie digitale Objektsammlungen
- Formatbeschreibungen der OAI- und SRU-Schnittstellen
- Erklärvideos
- Online Coding Tutorials
- Praktische Anwendungen und Codingbeispiele
- zentrale Ansprechpartner*innen

Ausflug ins DNBLab

Die Deutsche Nationalbibliothek bietet ihre bibliografischen Daten und mehrere digitale Objektsammlungen frei zugänglich an. Als zentraler Anlaufpunkt für die Präsentation, den Zugriff und die Nachnutzung der digitalen Ressourcen bietet das DNBLab einen Zugang zu unseren Daten, Objektdateien und Volltexten. Zugang erhalten Sie über verschiedene Schnittstellen und Downloads.

Jetzt SRU-Demoabfrage starten [↗](#)

Ein weiteres Angebot: Jupyter Notebook Tutorial für Einsteiger*innen **Wie können Daten mittels der SRU-Schnittstelle abgerufen werden?** [↗](#)

Eigene Metadaten sets können ohne Programmiercode über das DNB SRU Query Tool erstellt werden: Einfach Katalogbereich und gewünschtes Metadatenformat auswählen, Suchanfrage wie im Katalog eingeben und die ausgewählten Datensätze herunterladen!

DNB SRU Query Tool [↗](#)

Neben den im Rahmen der Metadatendienste regelmäßig aktualisierten Gesamtabzügen von Titel- und Normdaten stellen wir im DNBLab folgende (thematische) Datensets zur Verfügung.

Datensets			
Setname	Beschreibung	Anzeige im Katalog	Datensets mit Hyperlinks zu Objekten und/oder Bilddateien
Digitalisierte Inhaltsverzeichnisse	Mehr als 2,5 Millionen Inhaltsverzeichnisse von Büchern, die seit 1913 vor allem in Deutschland erschienen sind; zeitlicher Schwerpunkt sind die Jahre ab 1983 bis heute mit über 5 Millionen durchsuchbaren Einzelseiten		Download  (MARC21-xml)
Freie Netzpublikationen	Mehr als 750.000 Objekte ohne Zugriffsbeschränkung aus Deutschland	Sammlung im Katalog 	Download  (MARC21-xml)
Freie Online-Hochschulschriften	Mehr als 302.000 Online-Hochschulschriften ohne Zugriffsbeschränkung aus Deutschland	Sammlung im Katalog 	Download  (MARC21-xml)

OAI-Schnittstellen

Über OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) bieten wir zwei verschiedene Schnittstellen an.

Informationen zu der im Rahmen der Metadaten Dienste angebotenen OAI-Schnittstelle zur Abfrage von Metadaten und darüber verfügbaren Sets (Deutsche Nationalbibliografie, Normdaten und Zeitschriftendatenbank) finden Sie unter <https://www.dnb.de/oai>.

Beispielanfrage zur Ermittlung aller online frei verfügbaren Dissertationen der Sachgruppe „Sozialwissenschaften, Soziologie, Anthropologie“ aus einem bestimmten Zeitraum:

https://services.dnb.de/oai/repository?verb=ListRecords&from=2020-04-01T14:55:00Z&until=2020-07-08T09:54:59Z&metadataPrefix=oai_dc&set=dnb-all:online:dissertations:sg300


SRU-Schnittstelle

Über SRU (Search/Retrieve via URL) können Sie in den Daten der Deutschen Nationalbibliothek in Form von URLs oder über ein Computerprogramm (z.B. Python Bibliothek) suchen und eigene Datensets zusammenstellen. Anfragen werden wie im Katalog in der Retrievalsprache Contextual Query Language (**CQL**) formuliert und als Parameter an die URL angehängt.

Beispielanfrage zur Ermittlung aller online frei verfügbaren Titel mit dem Stichwort Klimawandel (tit=klimawandel and location=onlinefree):


```
https://services.dnb.de/sru/dnb?  
version=1.1&operation=searchRetrieve&query=tit3Dklimawandel\*+and  
+location%3Donlinefree&recordSchema=MARC21-xml
```

Die Ergebnisse können per XML in verschiedenen bibliografischen Formaten ausgeliefert werden. Weiterführende Informationen zu den SRU-Funktionen, Suchattributen, verfügbaren Katalogen und Formaten finden Sie unter **www.dnb.de/sru**.

Sehen Sie auch: Jupyter Notebook Tutorial **[Wie können Daten mittels der SRU-Schnittstelle abgerufen werden?](#)** 



BibSonomy

Die BibSonomy Genealogie ist Teil der sozialen Verschlagwortungsplattform BibSonomy und verfolgt das Ziel, eine freie Genealogie der deutschen Wissenschaft zu erstellen. Promovierte können sich sowohl den Betreuer*innen ihrer Doktorarbeit zuordnen als auch weitere wissenschaftliche Beziehungen und Publikationsdaten gemeinsam verwalten und abbilden. Daraus kann insbesondere ein Stammbaum der Forschung an deutschen Universitäten erstellt werden. Grundlage der Daten ist der Katalog der Deutschen Nationalbibliothek.

[BibSonomy](#) 

Kinder- und Jugendliteratur-Bot

Der Kinder- und Jugendliteratur-Bot (KJL-Bot) sammelt deutschsprachige Neuerscheinungen ausgewählter Verlage aus dem Bereich Kinder- und Jugendliteratur und führt sie strukturiert und durchsuchbar auf einer Website zusammen. Datenbasis bildet der Katalog der Deutschen Nationalbibliothek.

[KJL-Bot](#)  [GitHub](#) 

Sortiere KJL-Veröffentlichungen nach:

Autor*in

Verlag

Titel

Erscheinungsort

Erscheinungsdatum

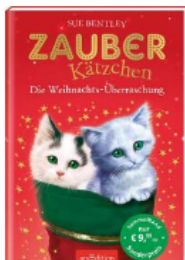
DJLP



KIMI



Neueste zuerst



Autor(in): Bentley, Sue

Titel: ZauberKätzchen – Die Weihnachts-Überraschung

Verlag: arsEdition

Erscheinungsort: München

Erscheinungsdatum: Oktober 2023

Schlagwörter: -

[Link zur DNB](#)



Autor(in): Wagner, Maja

Titel: WOW! Das Metallic-Stickerbuch – Winterzeit

Verlag: arsEdition

Erscheinungsort: München

Erscheinungsdatum: Oktober 2023

Schlagwörter: -

[Link zur DNB](#)

Code-Beispiele und Jupyter Notebooks


Liebe und Tod in der Deutschen Nationalbibliothek

Der Katalog der Deutschen Nationalbibliothek wird zum Forschungsobjekt. In dieser Untersuchung wurde ein Framework entwickelt, um bestimmte Aspekte des Katalogs anhand der Beispiele von Romanen zu untersuchen.

[Liebe und Tod in der Nationalbibliothek](#), [GitHub](#) 

Auf R basierte Schnittstellenabfrage der DNB-Daten

Beispiel einer automatisierten Abfrage der SRU Schnittstelle der Deutschen Nationalbibliothek mittels Programmiersprache R.

[GitHub](#) 

Code-Beispiele und Jupyter Notebooks

Wie können Daten mittels der OAI-Schnittstelle abgerufen werden?



Wie können Daten mittels der SRU-Schnittstelle abgerufen werden?



Wie können MARC-Titel analysiert werden?



Wie können freie digitale Objekte heruntergeladen werden?



Wie können digitalisierte Inhaltsverzeichnisse abgerufen und analysiert werden?



Wie können Daten aufbereitet und visualisiert werden?



Wie können Audio-Dateien abgerufen und analysiert werden?



Wie können Daten mittels Open Refine aufbereitet und angereichert werden?



Handouts

Hier finden Sie nützliche Übersichten und Zusammenstellungen zu unseren Daten als Handouts zu verschiedenen Themen:



[Welcher Zugang ist für Sie der richtige?](#)

257kB, PDF - Datei ist nicht barrierefrei



[Welches Exportformat ist für Sie das richtige?](#)

249kB, PDF - Datei ist nicht barrierefrei



[Übersicht: Inhalte in MARC21-Feldern](#)

172kB, PDF - Datei ist nicht barrierefrei



[Download von Objekten über Links in Datensets](#)

55kB, PDF - Datei ist nicht barrierefrei

Ausblick

- Neukonzeption freien digitalen Objektsammlungen
- weitere Online Coding Tutorials & Erklärvideos
- Einführungsveranstaltungen mit vertiefenden Workshops
- Workshops und Lehrkooperationen

Vielen Dank für Ihre Aufmerksamkeit!

Weiter geht's mit ...

... morgen mit Workshops 09:00 bis 13:00 Uhr