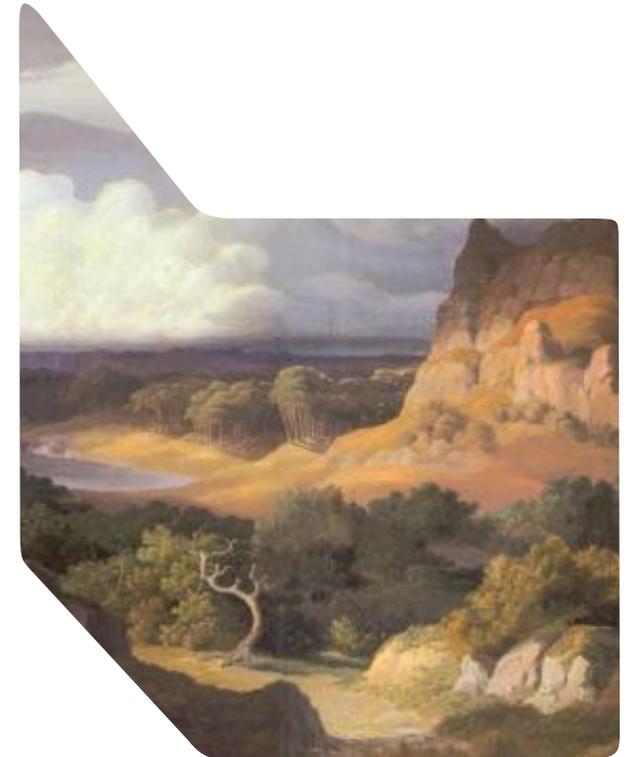


Daten öffnen, Forschung fördern

Initiativen mit denen die Zentralbibliothek Zürich die Digital Humanities unterstützt

Kolloquium der Universitätsbibliothek Bielefeld, 20.12.2024

Dr. Elias Kreyenbühl, Data Librarian am ZB-Lab



Inhalt

1. Kontexte I
2. Bildersammlung mit IIIF
3. Datenexport aus swisscollections.ch
4. Katalogdaten exportieren und analysieren mit Jupyter Notebooks
5. Kontexte II: Strategie UB+ZB, Produktportfolio

1 Kontexte

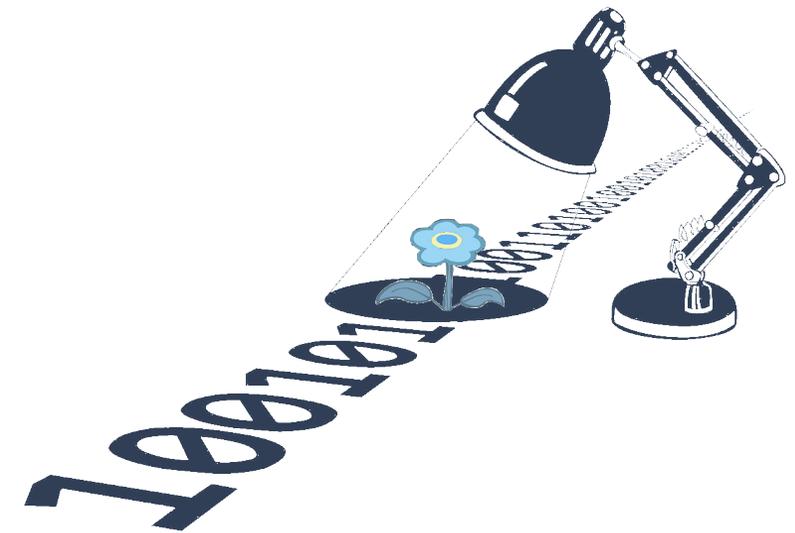
Kontext 1: Digital Humanities + Open Science

- Methodische Veränderung Richtung datenbasierter Auswertung, auch in den Geisteswissenschaften
- Neue Fragestellungen, die grosse Datenbestände in Betracht ziehen.
- zunehmendes Interesse an unseren Beständen als Daten
- Bibliotheksangebote noch stark auf "Suche mit zwei Augäpfeln" ausgerichtet

Kontext 2: ZB-Lab



Seit 2020. Mission ZB-Daten besser nutzbar zu machen.
Vermittlung von Datenkompetenz, Projektbegleitung, Prototypen



2 Bildersammlung mit DetektIIF

IIF

- Ist ein Standard, wie Bilder von Kulturerbeinstitutionen im Internet publiziert werden
- IIF basiert auf Linked Data und erlaubt eine direkte Nachnutzung
- Viele tolle Anwendungen, u.a. Annotation von Bildern
- Community & Open Source
- Potential für geisteswissenschaftliche Forschung

Use Case mit IIF

- Analog zum Gang in mehrere Archive und Sammlungen
- Browser Extension als Begleiter auf Archivreise
- Erstellen einer Sammlung
- Speichern und teilen der Sammlung
- Wiederverwendung & Daten werden vom Ursprung her geladen

Screencast: so funktioniert der DetektIIF



<https://youtu.be/21kxoJxbLV8?si=6D-ZrO8AD1UMWeBU>

Fazit

- Auf Basis von Open Source eine Open Source Anwendung erstellt
- Viele Schulungen und Konferenzen
- Adoption ist gering
- Evtl. zu sehr vom Technischen her gedacht oder
- Arbeitskultur der Zielgruppe falsch eingeschätzt

3 Datenexport aus swisscollections.ch

ZB-Lab Prototyp «OneAPI» 2022

- Zerklüftete Datenlandschaft. Systeme sind nicht miteinander verbunden
- Bibliothek hat gute und wertvolle Daten
- ZB-Daten wenig genutzt, weil schwierig
- Verbesserung Accessibility und Interoperability

Nutzerzentriertes Vorgehen



ZB-Lab «OneAPI»

2022

- einfacher Bulk-Download grosser Mengen Text-, Bild- und Metadaten
- Datenexploration aufgrund hochwertiger Metadaten
- Übersicht über Ergebnisse in CSV-Liste
- Zusätzliches Filtern anhand eigener Kriterien
- Datendownload per Script / API

ZB-Lab «OneAPI»

2022

- Kein «ready made», keine vorgefertigten Datensets sondern dynamische Erstellung über eine API
- Fokus auf Korpusbau, daher
- Explorieren der Daten als Voraussetzung
- Aber kein zusätzliches Frontend (!)
- Prototyp auf Swisscovery Sackgasse

ZB-Lab «OneAPI»

Hände Lavater

swisscovery (ohne CDI)



Erweiterte S...

Melden Sie sich für Bestellungen und weitere Dienstleistungen an Anmelden

VERWERFEN

0 ausgewählt SEITE 1 29 Ergebnisse



Download all

CSV Upload

1



ARCHIVMATERIAL

[20 Hände]

Lavater, Johann Caspar (Zusammenstellender)

ca. 1760 - 1801 [Handschriftliche Annotation datiert 13.08.1796]

Bestand prüfen >



alma99117227518605508

[Pnx](#)

[source record](#)

[OAI IZ](#)

[SRU IZ](#)

2



ARCHIVMATERIAL

90 [9 Hände]

Lavater, Johann Caspar (Zusammenstellender)

ca. 1760 - 1801

Bestand prüfen >



alma99117227298005508

[Pnx](#)

[source record](#)

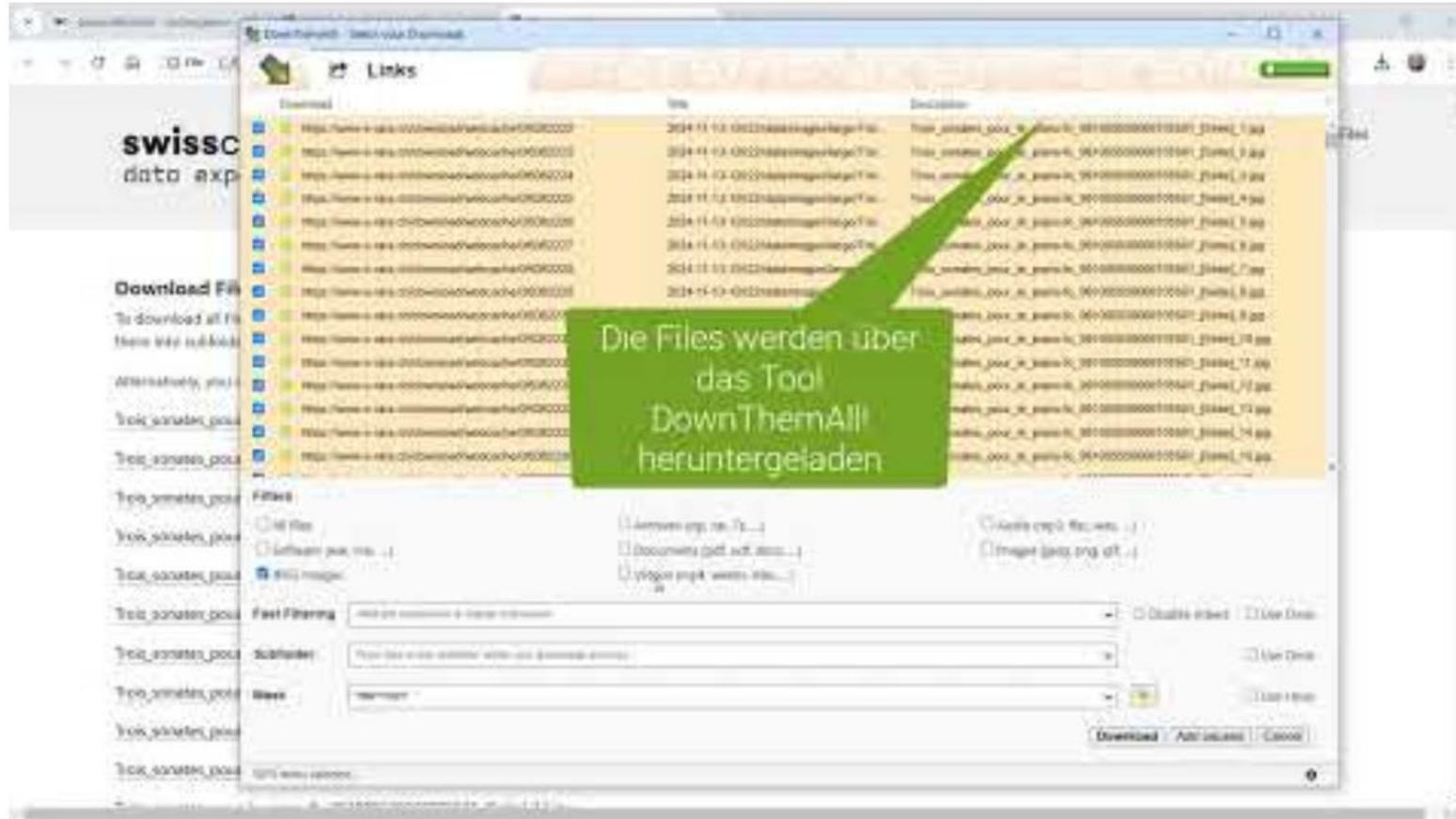
[OAI IZ](#)

[SRU IZ](#)

Produktiver Service in swisscollections.ch

- Metakatalog für historische und moderne Sammlungen in CH-Archiven und Bibliotheken.
- Handschriften, Archivbestände, Nachlässe, alte Drucke, Musikalien, Bildbestände und Karten.
- Verein bestehend aus 14 Archiven und Bibliotheken.
- Entwicklung und Betrieb durch die UB Basel.

Screencast: so funktioniert der Datenexport



<https://www.youtube.com/watch?v=k3tJyyYb-rg&vq=hd1080>

Fazit

- Glücksfall: vom Prototyp zur skalierten, betriebsfähigen Lösung
- (noch) keine Einschränkung der Treffermenge >10k dauert
- 30 Downloads/Tag
- «Damit hätte ich mir eine Menge Zeit gespart»
- Kleine und Mittlere Bibliotheken profitieren
- Es braucht den Austausch mit Nutzenden, mit anderen Bibliotheken
- Library Labs

4 Jupyter Notebooks

Katalogdaten beziehen

Linda Samsinger, Abteilung Erwerbung und Datamanagement

- Mit erweiterter Suche aus CH-Verbundkatalog SLSP
- Herunterladen als XLS
- Inhaltsverzeichnisse
- Mit visueller Datenauswertung
- Frei anpassbar

<https://gitlab.uzh.ch/zentralbibliothek-zuerich/zb-suche-katalog>

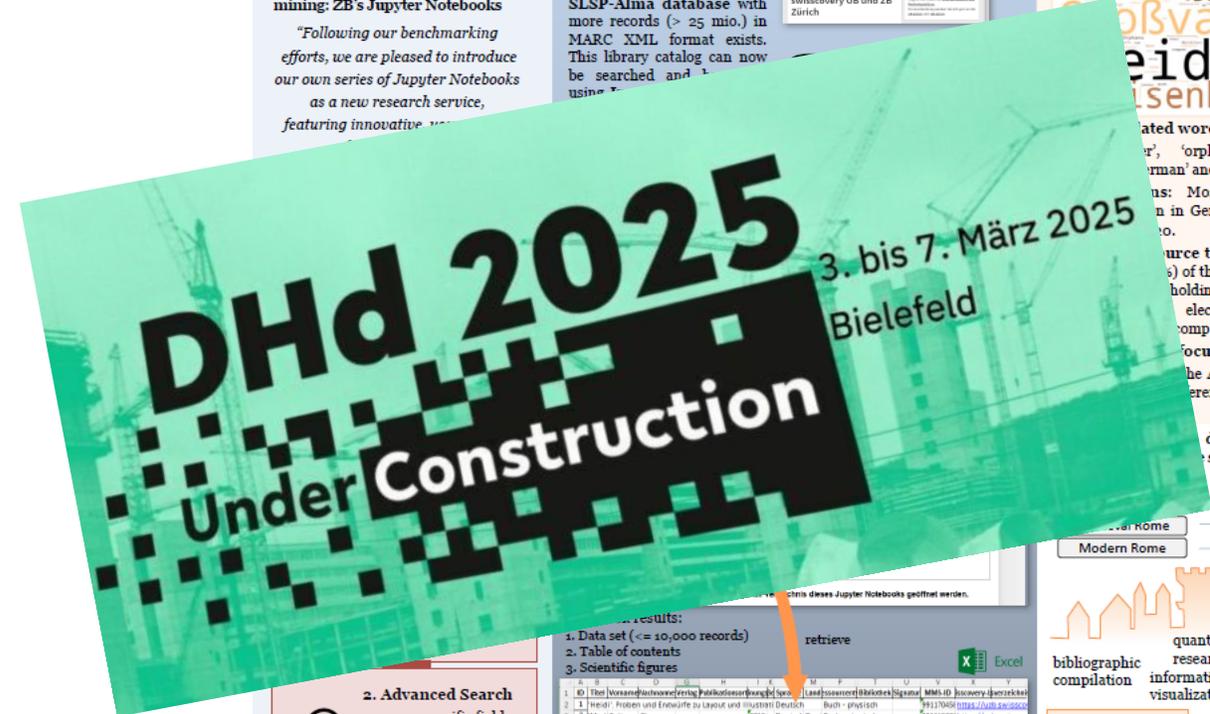
KoLibri - Leveraging Zurich Zentralbibliothek's Jupyter Notebooks for Metadata Retrieval and Analysis from Alma

Making library data available for researchers.

Linda Samsinger  ORCID: 0009-0004-6133-8790
Bielefeld | 3-7 March 2025

 digital humanities im deutschsprachigen raum

 Zentralbibliothek Zürich



Research | Literature Review
Digital Humanities | Data Mining
Library Catalog Search Methods
Library Management

① **New data-driven tools for data mining: ZB's Jupyter Notebooks**
"Following our benchmarking efforts, we are pleased to introduce our own series of Jupyter Notebooks as a new research service, featuring innovative..."

② **Reaping the crops off a Swiss-wide library catalog**
Typically, a library user searches the Swisscovery website (> 6 mio. records) for library data. Yet, another SLSP-Alma database with more records (> 25 mio.) in MARC XML format exists. This library catalog can now be searched and analyzed using Jupyter Notebooks.

③ **A search prompt example: Johanna Spyri's Heidi for data analysis**
Suchbegriff(e):
Freundschaft Deutsch Großvater Heidi Almenkind Waisenkind
related words: 'orphan', 'alpine', 'friendship'.
Mostly, physical in German between
source types: A small of the media items holdings consists of electronic audio computer files.
focus: the Alps, however, references to Japan

data sets for sky
Data sets
Modern Rome
quantitative research corpus building visualization
bibliographic information building visualization
Try it out:
Key benefits
✓ No login
✓ Open-source
✓ Transparency
✓ Usability vs. logic

2. Advanced Search
• query specific fields in the catalog for more targeted results

3. Data Analysis
• analyze the search results statistically
• create bar and pie charts, word clouds or world maps

4.-6. Enriched JNs
• add out-of-the-box GND/Wikidata data fields
• customize data fields

1. Data set (<= 10,000 records)
2. Table of contents
3. Scientific figures

retrieve
analyze

KoLibri is feeding on book flower nectar!
References: <https://data.zb.uzh.ch>

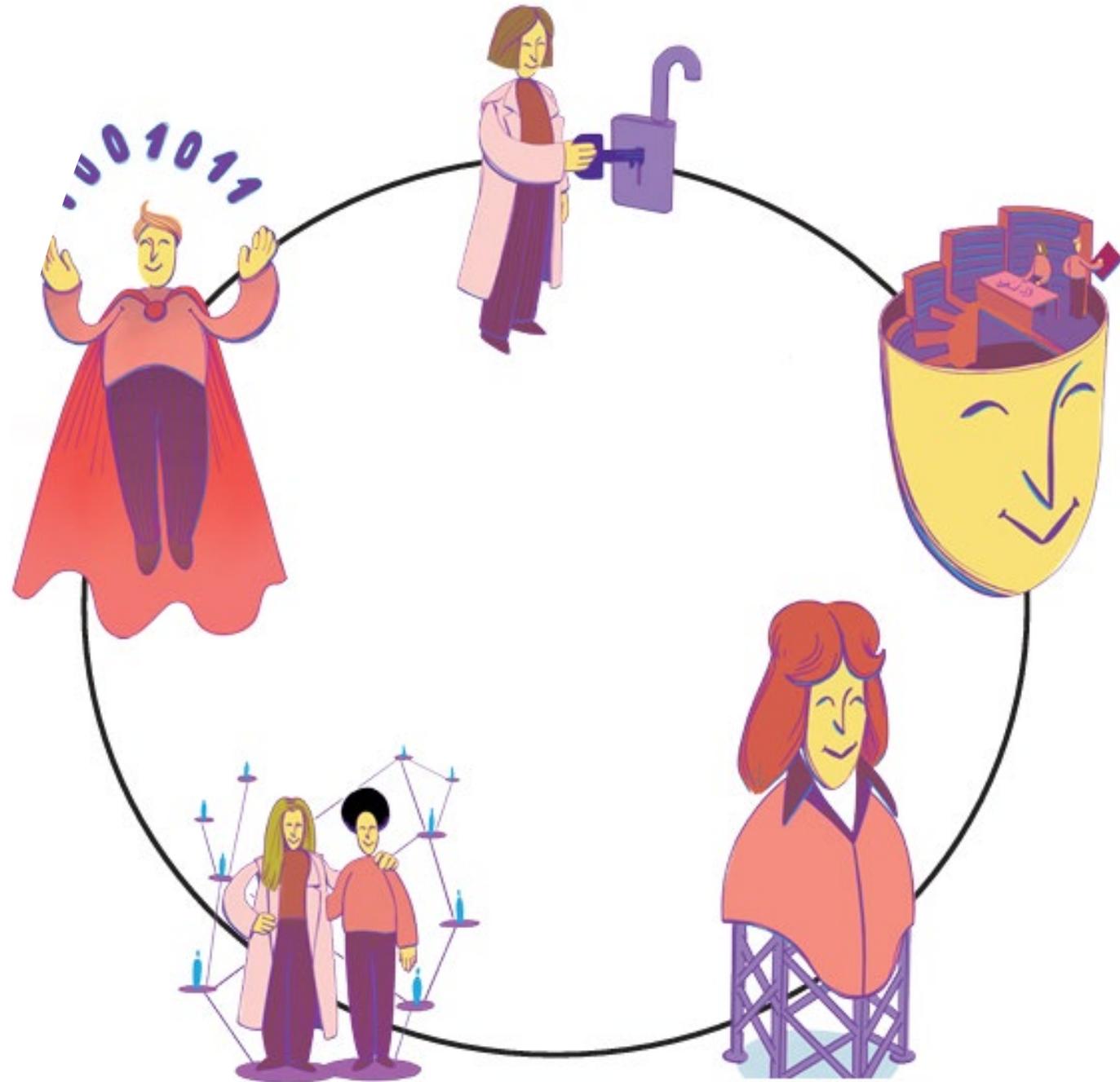
Contact Us:
Bibliotheksinformatik
Zurich Zentralbibliothek
bibliotheksinformatik@zb.uzh.ch
www.zb.uzh.ch

5. Kontexte II: Strategie UB+ZB, Produktportfolio

Gemeinsame Strategie UB und ZB 2024-27

5 Stossrichtungen

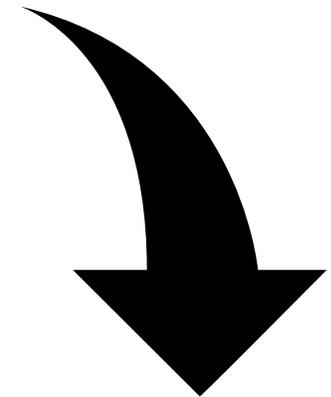
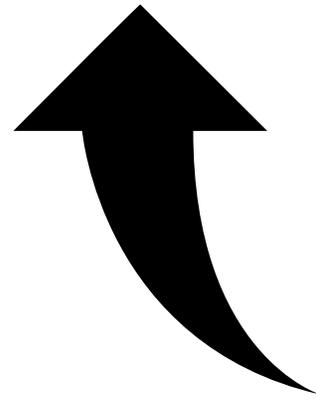
- Wir unterstützen Open Science
- Wir vernetzen die Gesellschaft und Wissenschaft
- Wir fördern die digitale Kompetenz
- Wir denken Räume neu
- Wir entwickeln unsere Mitarbeitende und Strukturen



Strategie: Open Science unterstützen

Produktfeld7: Forschungsunterstützung

Team Prozessunterstützung



Kerngruppe Forschungsservices
Abteilungen und Mitarbeitende

Aktuelle und künftige Forschungsservices

- Publikation von Datensets & Schnittstellen für Self Service
- Bereitstellen von spezialisierten Tools für die Forschung
- Vermittlungsangebote für DH: Praxis Atelier und Summer School
- Willy Bretscher-Fellowship; 12 Mte für DH Projekte und 20. Jh.
- Services für Digitale Editionen



swisscollections
swisscollections.ch

ZB-Lab

lab@zb.uzh.ch

