



h_da

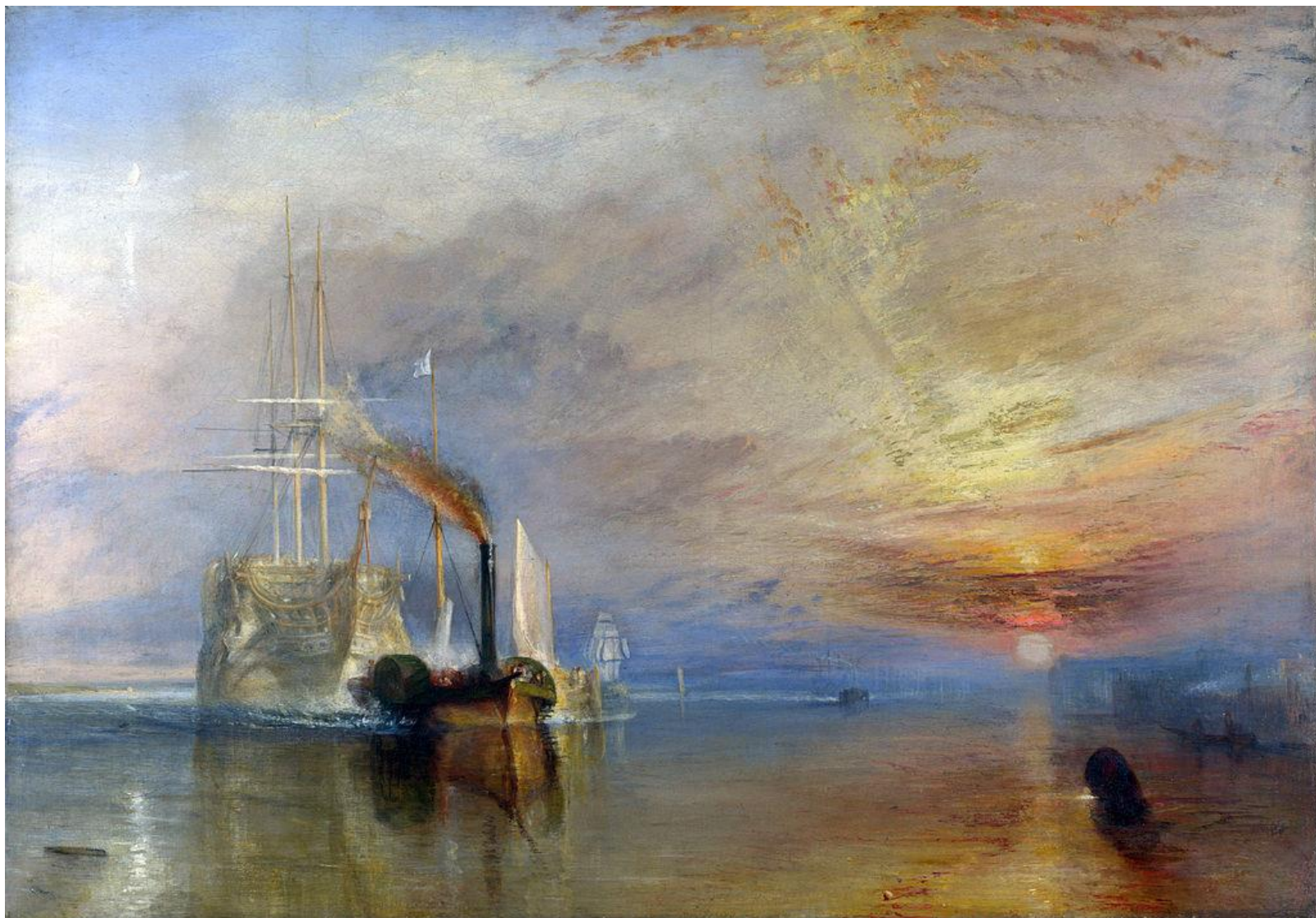
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

Open Data ... Open History? – Die Bedeutung historischer Forschungsdaten im digitalen Zeitalter

Stefan Schmunk - Hochschule Darmstadt

International Open Access Week – Staats- und Universitätsbibliothek Bremen

24. Oktober 2025



The Fighting Temeraire tugged to her last berth to be broken up, 1838. William Turner, National Gallery, London.
CC-0

„Digitalisierung“ als Phänomen einer Gleichzeitigkeit des Ungleichzeitigen

Quelle: https://www.youtube.com/watch?v=LcgG_E9gQJM

„Digitalisierung“ als disruptiver Prozess des kulturellen Erbes



Der Apple-CEO Steve Jobs präsentiert das erste iPhone in San Francisco. (Quelle: Kimberly White/Reuters)



<https://www.spiegel.de/netzwelt/web/google-plus-dienst-macht-am-2-april-endgueltig-dicht-a-1250986.html>



<https://www.zeit.de/2025/37/datenloeschung-usa-archivierung-wissenschaft-datenrettung>



© dapd - "Arabischer Frühling" 2011 -
<https://taz.de/Arabische-Revolution-und-soziale-Netzwerke/I5106835/>

Auswirkungen der Digitalität in den letzten 40 Jahren ...

1. Noch nie in der Menschheitsgeschichte wurden in solch einer kurzen Zeit so viele Daten erzeugt.
 2. Soziale und kulturelle Interaktionen finden zunehmend in digitalen Räumen statt, die von kommerziellen Plattformen betrieben werden, und dies unter der dauerhaften Aufgabe der individuellen Nutzungs- und Verwertungsrechte.
 3. Das digitale kulturelle Erbe weist eine immense materielle Flüchtigkeit und Zeitgebundenheit auf, so dass ein Großteil dessen von zukünftigen Generationen nicht mehr aufgefunden und genutzt werden kann.
 4. Die Komplexität und Volatilität unserer Gesellschaften und Lebensweisen verändern sich im 21. Jahrhundert mit einer hohen politischen Dynamik und einer Zunahme an Unsicherheiten und Mehrdeutigkeiten. Nach einer kurzen Phase der Demokratisierung in den 1990er Jahren erfolgte eine Zunahme von Instabilität (John Mearsheimer 1990).
- Das 21. Jahrhundert wird sich für zukünftige Generationen aus einer Perspektive der Überlieferung für die Geschichtswissenschaft zu einem **Dark Age** und zugleich aus politischer, gesellschaftlicher und technologischer Perspektive zu einer **digitalen Sattelzeit** entwickeln.

Quellen & Forschungsdaten

Digitalisate / Retrodigitalisate sind in ihrer originären digitalen Form keine Quellen, sondern Daten, die erst durch Kontextualisierung zu Quellen werden.

„Quellen nennen wir alle Texte, Gegenstände und Tatsachen, aus denen Kenntnis der Vergangenheit gewonnen werden kann“

(Paul Kirn 1986, S.29)

Forschungsdaten

„Daten, unabhängig ihrer Provenienz, werden im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage zu digitalen Forschungsdaten, sobald sie gesammelt, beschrieben, ausgewertet und/oder erzeugt und in **maschinenlesbarer Form** zum Zwecke der Nachvollziehbarkeit von Forschungsergebnissen sowie zur Archivierung, Zitierbarkeit und Weiterverarbeitung aufbewahrt werden. Eine Aggregation von Forschungsdaten ($n > 1$) ist in diesem Sinne eine Wissenschaftliche Sammlung.“

Quelle: Oltersdorf/Schmunk: [Von Forschungsdaten und Sammlungen](#).

Datenmodellierungslayer textbasierter Daten

- I. Digitale, **maschinenlesbare Metadaten** ohne Digitalisate
- II. Digitale Repräsentation von analogen Daten – **Bild-Digitalisate**
- III. **Volltexte unstrukturiert** bzw. nicht standardisierte Strukturierung

- A. Roh-OCR
- B. optimierte OCR (> 99 % Zeichengenauigkeit)
- C. Manuelle Erfassung (> 99,97 % Zeichengenauigkeit)
- D. Volltexte (per OCR/OLR und/oder manuell erzeugt)

Aktueller Stand:
Erst 50-60% von VD16,
VD17 und VD18 sind
bislang als Bild-
Digitalisate vorhanden

- IV. **Volltexte mit standardisierter Strukturierung** (maschinell weiterverarbeitbar)
- V. **Volltexte angereichert**, z.B. mit Entitäten und komplexen Strukturmerkmalen

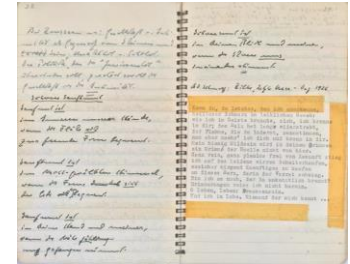
- A. angereichert mit Normdaten (z.B. GND, VIAF, TNG, ...)
- B. Annotation
- C. Kollation
- D. Digitale Edition

Forschungsdaten

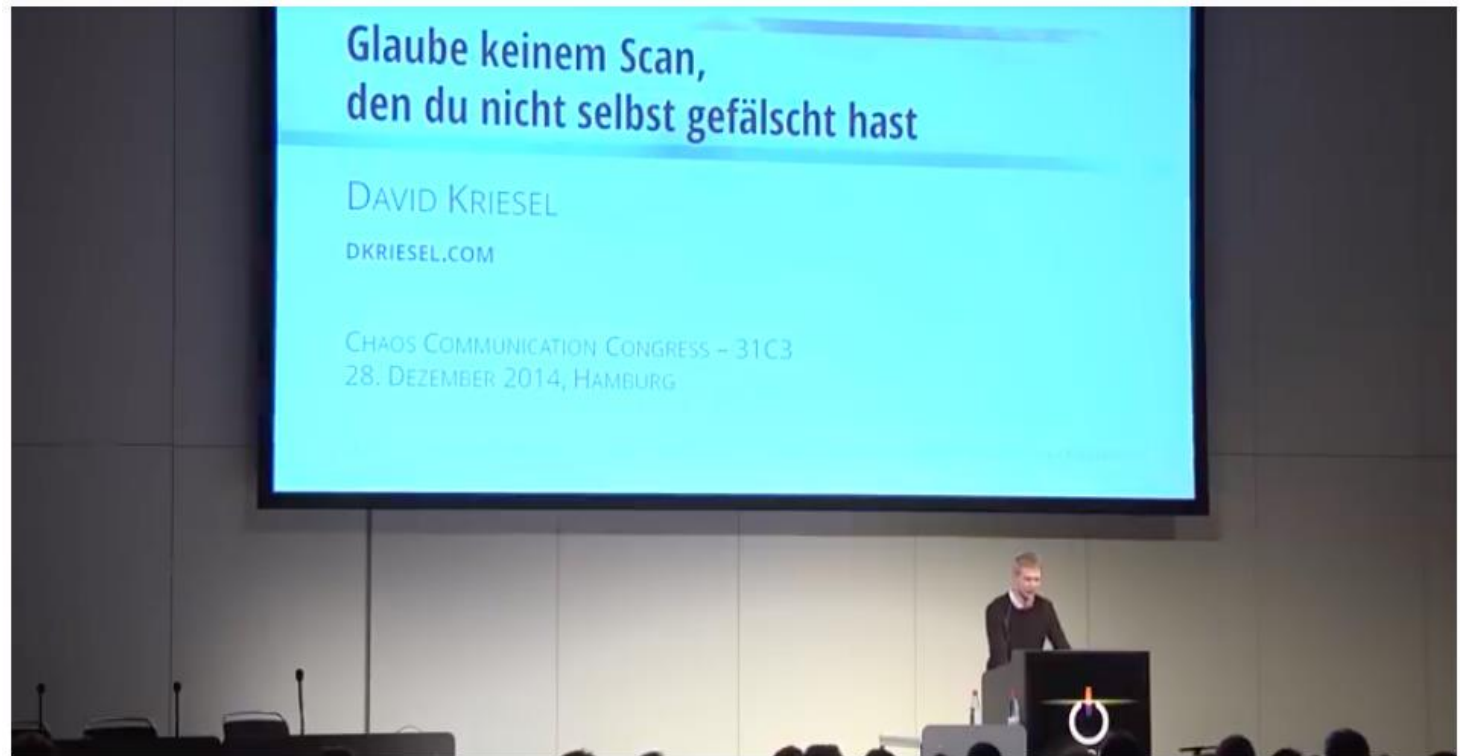
Quelle: Klaffki/Schmunk/Stäcker: Digitalisierung von Kulturgut: [urn:nbn:de:gbv:7-dariah-2018-1-3](https://nbn-resolving.org/urn:nbn:de:gbv:7-dariah-2018-1-3)

Forschungsdaten in den Geschichtswissenschaften

- Metadaten, bibliographische Daten, Kataloge
 - Bild-Digitalisate analoger Daten und Quellen
 - Digitale Objekte
 - Volltexte und Transkriptionen
 - Angereicherte Volltexte
 - Bilder, Filme, Musik, Noten
 - Normdaten, kontrollierte Vokabularien, Ontologien
 - Biographische Daten
 - Annotationen, Kollationen, Kommentierungen
 - Exzerpte
 - etc.
- 

[illegible]

Xerox Scanning Bug: „Traue keinem Scan, den Du nicht selbst gefälscht hast ...“



Bilddigitalisate sind eben nicht eine identische
Abbildung eines analogen Objektes

Quellen: <https://www.youtube.com/watch?v=7FegF1-Z1g0>

http://www.dkriesel.com/blog/2013/0802_xerox-workcentres_are_switching_written_numbers_when_scanning

NFDI4Memory

... und mehr als 80 nationale & internationale Partner



Aufgaben und Ziele

Digitale Transformation in den historischen Geisteswissenschaften gestalten

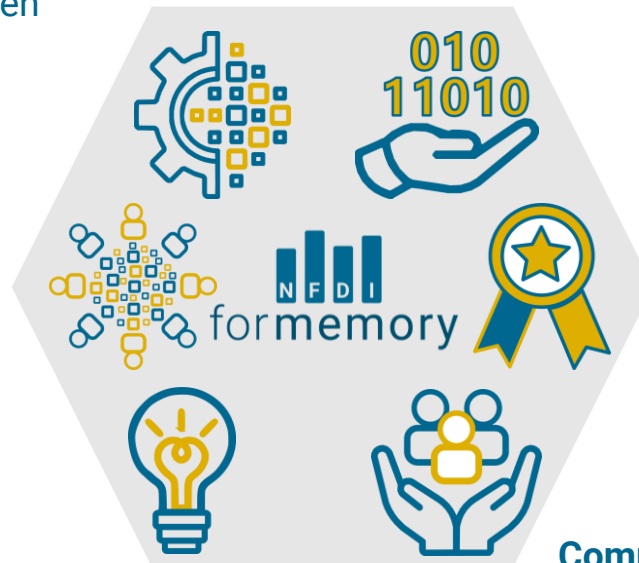
Nachhaltige Sicherung & Archivierung relevanter Forschungsdaten

Datenkultur in historischen Disziplinen fördern

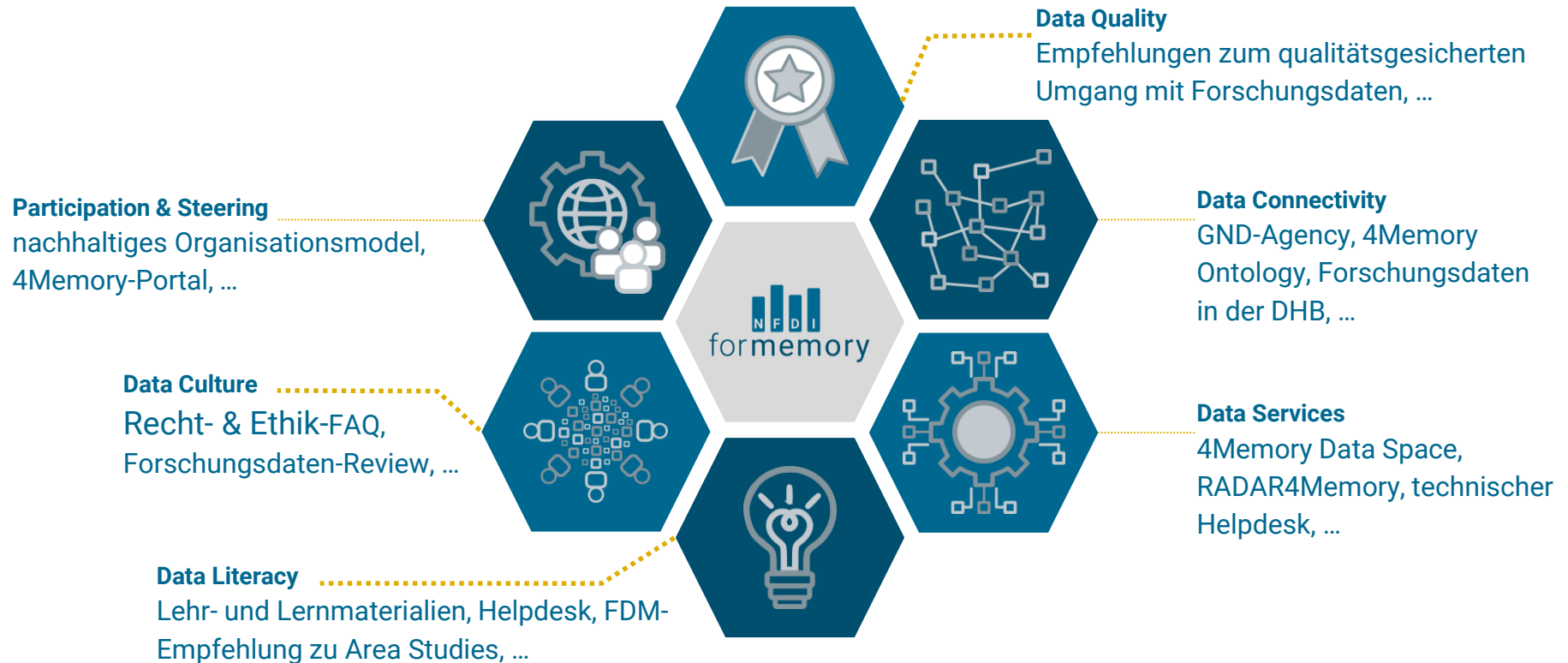
Qualitätsstandards für Digitalisierung & Datafizierung historischer Quellen etablieren

Data Literacy-Kompetenzen verbreiten & vertiefen

Community unterstützen bei digitalen Tools & Forschungsdatenmanagement

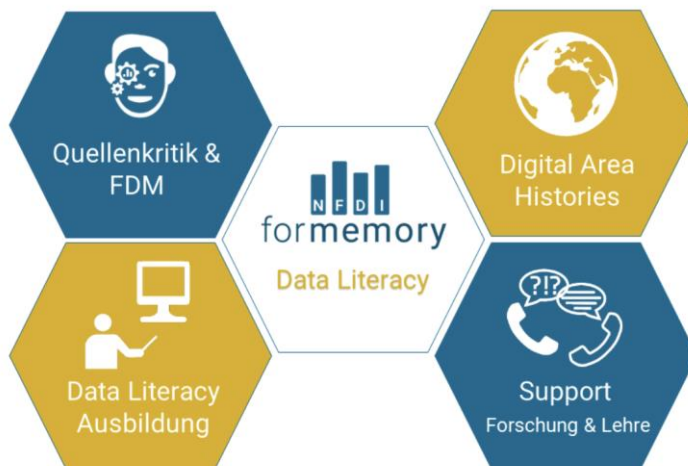


Arbeitsschwerpunkte



Task Area: Data Literacy

- Vermittlung von Datenkompetenzen in Forschung und Lehre in den historischen Disziplinen Unterstützung der Community beim digitalen Wandel
- Entwicklung von Services and Lehrangeboten zu digitalen Datenkompetenzen



Weitere Aktivitäten

Arbeitsgruppe Data Literacy in der Lehre	+
RDMTraining4NFDI	+
Arbeitsgruppe FDM-Helpdesks Netzwerk	+
OER.net-Gruppe	+
Unterarbeitsgruppe Schulungen/Fortbildungen	+
Arbeitsgruppe "FDM-Beschreibungsmodell"	+

<https://4memory.de/ueber-4memory/task-areas/data-literacy/>

Service HISTOCAT

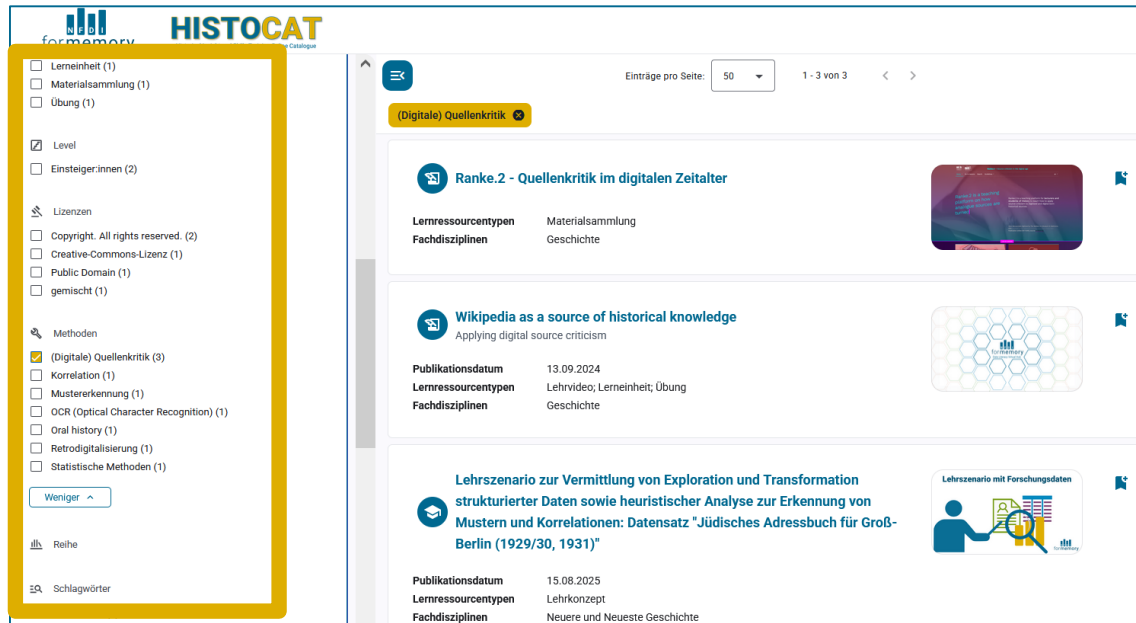
Der Historical Insight and Skills Training Online Catalogue (HISTOCAT) ist eine kuratierte Sammlung von Lehr- und Lernmaterialien zum Erwerb und zur Vermittlung von Datenkompetenzen in den historisch arbeitenden Disziplinen.



<https://histocat.uni-trier.de/#/>

Materialsammlung: HISTOCAT

Beispiel: Materialien zur Digitalen Quellenkritik (Methoden)

The screenshot shows the HISTOCAT web interface. On the left, a sidebar contains filter categories: 'Lerneinheit' (1), 'Materialsammlung' (1), 'Übung' (1), 'Level' (with 'Einsteiger:innen' at 2), 'Lizenzen' (Copyright, Creative Commons, Public Domain, and gemischt), 'Methoden' (with '(Digitale) Quellenkritik' at 3, and other methods like Korrelation, Mustererkennung, OCR, Oral history, Retrodigitalisierung, and Statistische Methoden), 'Reihe', and 'Schlagwörter'. The main area displays search results for '(Digitale) Quellenkritik'. Three results are visible: 'Ranke.2 - Quellenkritik im digitalen Zeitalter' (Materialsammlung, Fachdisziplinen: Geschichte), 'Wikipedia as a source of historical knowledge' (Applying digital source criticism, Publikationsdatum: 13.09.2024, Lernressourcentypen: Lehrvideo; Lerneinheit; Übung, Fachdisziplinen: Geschichte), and 'LehrszENARIO zur Vermittlung von Exploration und Transformation strukturierter Daten sowie heuristischer Analyse zur Erkennung von Mustern und Korrelationen: Datensatz "Jüdisches Adressbuch für Groß-Berlin (1929/30, 1931)"' (Publikationsdatum: 15.08.2025, Lernressourcentypen: Lehrkonzept, Fachdisziplinen: Neuere und Neueste Geschichte).

Vokabular und
Metadaten für
Erschließung
(facettierte
Filterung)

Open Data ... Open History?

1. Open Data ist Voraussetzung für Open History
→ FAIR und CARE als Voraussetzung
2. Forschungsdaten werden erst durch Kontextualisierung zu geschichtswissenschaftlich nutzbaren Quellen
→ Digitalisate sind neue Repräsentationen von Daten
3. Open History benötigt einen methodischen Rahmen
→ Digitale Quellenkritik als Gegenstand in Forschung und Lehre

Herzlichen Dank für Ihre Aufmerksamkeit

Prof. Dr. Stefan Schmunk

Hochschule Darmstadt – University of Applied Sciences

stefan.schmunk@h-da.de