

SODa Kick-off

Johannes Schäffer, Martin Stricker, Louise Tharandt
11. April 2024

SODa Kick-off

- Basics zum Projekt, Wer sind wir?
- Umfrage: Wer sind Sie?
- Datenkompetenzen für wissenschaftlichen Objektsammlungen
- Geplantes Programm des Datenkompetenzzentrums
- Diskussion: Erwartungen an das Projekt
- Weitere Kick-offs & SODa-Termine

SODa

Sammlungen, Objekte, Datenkompetenzen

Das Verbundprojekt zum Aufbau eines
Datenkompetenzzentrums für wissenschaftliche
Universitätssammlungen

Aufbau: November 2023 - November 2026

Danach: Verstetigung (🤖)

<https://sammlungen.io/>



GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung



**Finanziert von der
Europäischen Union**
NextGenerationEU

Wissenschaft

15.11.2023

Aufbau von Datenkompetenzzentren an Hochschulen und Forschungseinrichtungen

Die Arbeit mit Forschungsdaten gewinnt immer weiter an Relevanz. Um das darin liegende Potenzial besser nutzbar zu machen, fördert das BMBF zunächst die Konzeption und später den Aufbau von Datenkompetenzzentren in der Wissenschaft, an denen Unterstützungs- und Vernetzungsangebote für Wissenschaftlerinnen und Wissenschaftler und weitere Akteure geschaffen werden. Alle 11 Projekte der Umsetzungsphase sind nun gestartet.

<https://www.bildung-forschung.digital/>



MEHR ZU:

→ [Zurück zur Übersicht](#)

Open Access

Urheberrecht

Weizenbaum-Institut

Forschungsdaten

Datenkompetenzen

Transfer-Hub Datenkompetenzen

Transfer-Hub Datenkompetenzen Newsletter

Roadmap Datenkompetenzen und Datenkultur

Datenkompetenzen für den wissenschaftlichen Nachwuchs

Datenkompetenzzentren für die Wissenschaft

Forschungsdatenmanagement an Fachhochschulen



FINANZIERT VON DER EUROPÄISCHEN
UNION – NEXTGENERATIONEU

Wer ist SODa?

- Projektleitung
- Partner*innen
- Programmentwicklung
- Fachexpertise
- Infrastrukturentwicklung

<https://sammlungen.io/team>



Wer sind Sie?

Live-Umfrage:

<https://sammlungen.io/pingo>



Datenkompetenzen

für die Arbeit in und mit wissenschaftlichen Objektsammlungen

Kompetenzen = Ich kann's!

Daten & wiss. Objektsammlungen



Daten & wiss. Objektsammlungen



Objekte =
(uncodierte) Daten

Jahr: 1880
Vierteljahr: April - Juni

Inventar-Nr.	Acten-Nr.	Katalog-Nr.	Gegenstand.	Provenienz.	Von wem erworben oder geschenkt?
772 773	177 178	344-113	Sammlung von Perlen de Chamak, Japan	Japan	Von dem Kaiser Geschenkt
788 789	181 182	344-113	Wappensteinen von Japan	Japan	Von dem Kaiser Geschenkt
811 812	183 184	344-113	Sammlung japanischer Perlen	Japan	Von dem Kaiser Geschenkt
813 814	185 186	344-113	Wappensteinen von Japan	Japan	Von dem Kaiser Geschenkt

analoge (Meta-)Daten



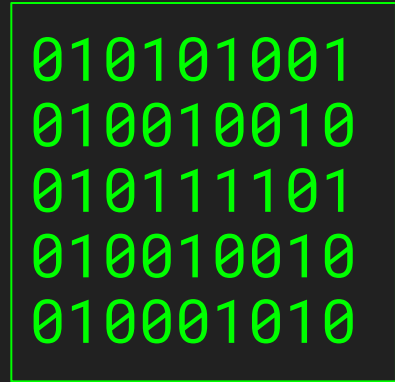
digitale Daten

Daten & wiss. Objektsammlungen



Jahr: 1880
Vierteljahr: April - Juni

Inventar-Nr.	Acten-Nr.	Katalog-Nr.	Gegenstand.	Provenienz.	Von wem erworben oder geschenkt?
7	272 279	177 344-118	Sammlung von Steinchen de Lamark, Japan	1880 Japan	Von Herrn v. Alphonse Tottor gebildet
8	708 210	107114 102288-29 2357 102288-29 102288-29 102288-29	Waffen des Okinawa Kamoyama	Japan Ok. von Ok. von Ok. von Ok. von	Von Herrn v. Alphonse Tottor gebildet
9	814 2679	102288-29 - 2790	Sammlung japanischer oder japanischer Gegenstände zur Topographie	Japan	Herrn v. Alphonse Tottor gebildet
10	247 276 117	102288-29 - 2790	2 Gegenstände mit Bergeln in Japan?	Japan	V. Alphonse Tottor gebildet V. Alphonse Tottor gebildet V. Alphonse Tottor gebildet



Objekte =
(uncodierte) Daten

analoge (Meta-)Daten

digitale Daten

OCR & ML

digitale Reproduktion

Decodierung & Erschließung der Kontexte

Daten & wiss. Objektsammlungen



Jahr: 1880
Vierteljahr: April - Juni

Inventar-Nr.	Acten-Nr.	Katalog-Nr.	Gegenstand.	Provenienz.	Von wem erworben oder geschenkt?
7	272 279	177 344-118	Sammlung von Stein de Lamak, Japan	1880 Japan	Von dem alpbach Tortosa gebildet
8	708 210	107114 102288-29 2357 102145-10 102177-10 102184	Waffen des Mann von Otaga	Japan 1880 1880 1880 1880	Von dem O. Thomas in Otaga 1880 1880 1880 1880
9	311 2679	102240 - 2790	Sammlung japanischer Papierarbeiten	Japan	1880 1880 1880 1880
10	243 276 117	102251 - 2758	2 Gegenstände mit Bergeln in Japan	1880 Japan	1880 1880 1880 1880



Objekte =
(uncodierte) Daten

analoge (Meta-)Daten

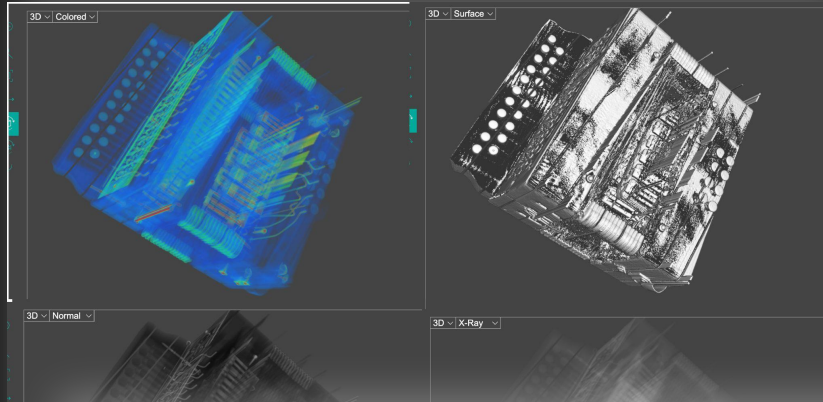
digitale Daten

OCR & ML

digitale Reproduktion

Decodierung & Erschließung der Kontexte

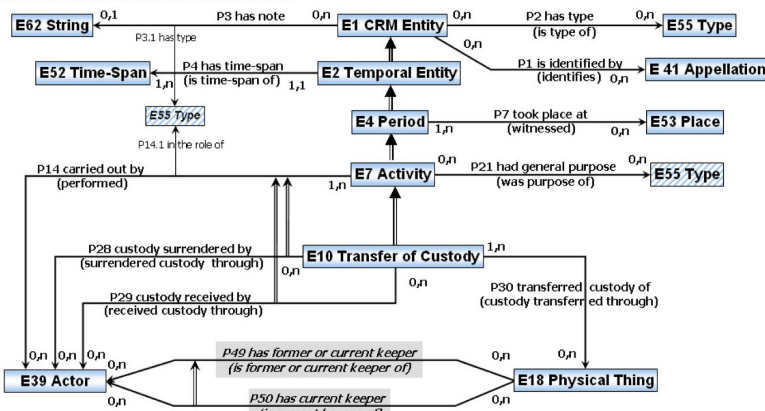
Digitale Datentypen wiss. Objektsammlungen



Digitale Reproduktionen

- 2D-Fotografien, 3D-Modelle, Audio, Video, AV, Animationen

OBJECT COLLECTION INFORMATION



Strukturierte Erschließungsdaten

- am oder vom Objekt erworben (Vermessen, Bestimmen, Vergleichen, Analysen ...)
- Kontext & Geschichte (Quellen, Sammlungskontexte, Normdaten/referenzierte Entities...)

Datenkompetenzen für wiss. Objektsammlungen

- Ich kann's
- ... mit den für wiss. Objektsammlungen spezifischen Datentypen
- ... für die Ziele und Zwecke der Arbeit mit wiss. Objektsammlungen an Universitäten und Hochschulen

Ziele und Zwecke der Arbeit mit wiss. Objektsammlungen an Universitäten und Hochschulen

Kuratorisch & Kustodial

- Sammlungen bewahren und entwickeln
- Sammlungen sichtbar und nutzbar machen

Objekte nutzen

- in Forschung
- in Lehre
- im Transfer (“Third Mission”, “Wissenschaftskommunikation”)

Ziele und Zwecke der Arbeit mit wiss. Objektsammlungen an Universitäten und Hochschulen

Kuratorisch & Kustodial

- Sammlungen bewahren und entwickeln
- Sammlungen sichtbar und nutzbar machen

=> **Kompetenzbereich: Datenlebenszyklus für Sammlungs- und Objektdaten**

Objekte nutzen

- in Forschung
- in Lehre
- im Transfer (“Third Mission”, “Wissenschaftskommunikation”)

=> **Kompetenzbereich: Forschung mit Objekt- und Sammlungsdaten**

Kompetenzbereich

Datenlebenszyklus für Sammlungs- und Objektdaten

Datenlebenszyklus für Sammlungs- und Objektdaten

Spezifik

Spezifik der Daten zu materiell-physischen Objekten:

- Spezifische Datentypen: Digitale Reproduktionen (“Digitalisate”) und strukturierte semantische Erschließungsdaten
- Daten sind unauflöslich mit dem materiell-physischen Objekt verbunden

Granularität & Vernetzung:

- Erschließungsdaten sind auf granularer Ebene mit vielen anderen Objekten und “Entities” (Akteur:innen, Orte, Ereignisse, ...) zu komplexen semantischen Netzen verknüpft
- Jedes Objekt sollte in seinen jeweiligen multiplen Bezügen granular “adressierbar” (auffindbar, auswertbar) sein

Datenlebenszyklus für Sammlungs- und Objektdaten

Spezifik

Kustodiale Verantwortung:

- Daten referenzieren ein materielles Objekt, welches häufig über lange Zeiträume bewahrt und genutzt wird
- Sammlungsverantwortliche und Kustod*innen verstehen sich oft auch verantwortlich für ein langfristiges Kuratieren der digitalen Daten zu den Objekten
- Insbesondere bedeutet dies auch Aktualisierungen der Daten aufgrund neuer Forschungen oder entsprechend neuer Standards

Datenlebenszyklus für Sammlungs- und Objektdaten

1. Planung von Erschließungs- und Digitalisierungsvorhaben; Konzeption, Konfiguration und Nutzung von Dateninfrastrukturen
2. Erzeugung von Objekt- und Sammlungsdaten
3. Gewährleistung von Datenqualität und Kompatibilität
4. Inhaltliche Aufbereitung und Anreicherung
5. Rechtliche und ethische Aspekte, Lizenzen
6. Langzeitverfügbarkeit nach FAIR-Prinzipien

Kompetenzbereich

Forschung mit Objekt- und Sammlungsdaten

Forschung mit Objekt- und Sammlungsdaten

1. Objekt- und Sammlungsgeschichte, Provenienzforschung und -dokumentation
2. Nutzen, Analysieren, Nachnutzen auf Basis von strukturierten Daten und insbesondere Graphen
3. Digitale Reproduktionen & 3D
4. Verfahren automatischer Datenanalyse (Machine Learning, Generative AI)
5. Konservierungs- und Restaurierungsdokumentation

Datenkompetenzen sind für uns ein Moving Target

Wir gehen von einer hohen Heterogenität und Dynamik in den angesprochenen Communities aus:

- beim Grad der Digitalisierung, Entwicklung digitaler Methoden, bei der Vernetzung und Standardisierung
- bei nachgefragten Datenkompetenzen

Deshalb:

- Wir setzen auf Impulse aus der Community, auf die Formulierung von Anforderungen und auf Ideen, Innovationen, Good Practice-Beispiele
- Wir fördern Vernetzung sowie Ideen- und Methodentransfer unter den Sammlungen und interdisziplinär von Fachcommunity zu Fachcommunity

Programm

Prinzipien

- Nutzer*innenorientiert
- praxis first
- bring your own data
- Lernort
- Forschungsort
- Vernetzungsort

Beratung

- Beratung und Begleitung von Erschließungs- und Forschungsvorhaben
- zentrale Adresse: soda@sammlungen.io
- Expert*innen für Kompetenzfelder: Erschließung & FDM, Digitale Provenienzforschung, Recht und Ethik, Maschinelles Lernen, 3D, Analyse strukturierter Daten
- Kooperationen fördern, Vernetzen von Akteur*innen

Basiskurs “Datenlebenszyklus”

- Format: Open Educational Resources und regelmäßige Veranstaltung
- Zielgruppe: Forschende und Sammlungsbetreuer*innen mit wenig Erfahrung in datengetriebener Arbeit
- Planung von Digitalisierungs-/Forschungsvorhaben
- Erschließung/Erzeugung von Objektdaten
- Sicherung von Datenqualität
- Recht und Ethik

Workshops und Tutorials

- primär online; ca. 3 mal jährlich in Präsenz
- Tutorials: Grundständiger Kompetenzaufbau
- Workshops: Vertiefung und Ausbau von Kompetenzen, aktuelle Probleme und Entwicklungen
- Beispiele
 - Softwarenutzung im Sammlungskontext
 - LIDO
 - 3D-Digitalisierung
 - Datenbereinigung mit OpenRefine
 - Datenanreicherung mit Normdaten
 - Maschinelles Lernen z.B. Bilderkennung, Texterkennung

Open Educational Resources

- Nachschlagewerk
- frei(e) zugängliche Informationsmaterialien zum Selbststudium
- aktuelle Standards und Good Practice
- Grundständiger Kompetenzaufbau

Semantic Coworking Space

- Virtuelle Arbeits- und Kommunikationsumgebung, in der kollaborativ und explorativ an (realen) Daten gearbeitet und geforscht werden kann
- => „bring your own data“
- kostenfreie Cloud-Lösung
- Tools zum Modellieren, Erschließen, Anreichern, Mapping, Analysieren und Auswerten von Daten
- Werkzeuge: WissKI, OpenRefine, Jupyter, WebProtegé
 - Weiterentwicklung und Integration von Werkzeugen
- Self-Service und Verwendung in SODa-Veranstaltungen

SODa Foren

- virtuelle monatliche Veranstaltungen
- Fokus auf ein spezifisches Thema/eine Herausforderung
- Werkstattberichte
- Präsentationen neuer Verfahren
- themenbezogene Impulse und Diskussionen

SODa Barcamp

- 14.-16. Oktober 2024 in Berlin
- Gestaltung durch die Teilnehmer*innen

Summer School

- 5 Tage intensiv-SODa
- 2025 in Berlin
- 2026 in Nürnberg/Erlangen
- für Nachwuchswissenschaftler*innen

Wissenschaftskommunikation

- Website sammlungen.io (noch vorläufig)
- Schaufenster der Möglichkeiten
- Blog
- Social Media

Vernetzung und Gremienarbeit

- Perspektive der Unisammlungen einbringen
- Weiterentwicklung von Standards
- Kooperationen fördern
- Netzwerk(e) universitärer Sammlungen
- NFDI 4Culture 4Objects 4Biodiversity 4Memory Text+
 - Sektion EduTrain
- Datenkompetenzzentren
 - WiNoDa
 - HERMES
 - QUADRIGA
- DHd e.V.

Diskussion und Erwartungen

- Etherpad: <https://ishtar.iwus.org/p/SODa-Kick-off>

Weitere Kick-offs & SODa-Termine

- 2.5. 14-16 Uhr Kick-off “Erschließung und sammlungsbezogenes Forschungsdatenmanagement”
- 22.5. 10-12 Uhr Kick-off “KI/Maschinelles Lernen”
- 26.6. 10-12 Uhr Kick-off “3D”
- es folgen: Ethik und Recht, Digitale Provenienzforschung, Konservierungs- und Restaurierungsdokumentation
- 14.-16. Oktober Barcamp in Berlin

Abbildungsnachweise

Folie 10:

- **Azurit.** Inventarnummer: GZG.HST.4.589, Sammlung: Geowissenschaftliches Museum, Akademisches Museum der Universität Göttingen, Lizenz: CC-BY-SA 4.0, Link: https://sammlungen.uni-goettingen.de/objekt/record_naniweb_366252/ (17. April 2024)
- **Jomon-Figur**, japanisches Fruchtbarkeitsidol (Nachbildung). Inventarnummer: K 493, Sammlung Heinz Kirchhoff der Universität Göttingen, Lizenz: CC-BY-SA 4.0, Link: https://sammlungen.uni-goettingen.de/objekt/record_kuniweb_365854/ (17. April 2024)
- Evakuierbare Glasglocke. Inventarnummer: H444, Sammlung: Physicalisches Cabinet, Lizenz: CC-BY-SA 4.0, Link: https://sammlungen.uni-goettingen.de/objekt/record_kuniweb_365854/ (17. April 2024)

Abbildungsnachweise

Folie 11:

- **Erwerbungsbuch**. Nachweiszeitraum der Zugänge 1880-1906, Inventarnummer: Inv. 3, Inventarverzeichnis I MV, S. 0013 [Ausschnitt], Ethnologisches Museum der Staatlichen Museen zu Berlin, Linz: gemeinfrei, Link:
https://storage.smb.museum/erwerbungsbaecher/EB_EMB-B_SLG_SN_0003-0003_LZ_1880-1906.pdf

Folie 14:

- **Handharmonika (CT-Scan)**. Inventarnummer: 4199, Sammlung: Museum für Musikinstrumente der Universität Leipzig, CT-Scan des Germanischen Nationalmuseums.
<https://musices.gnm.de/wisski/navigate/931/view>