



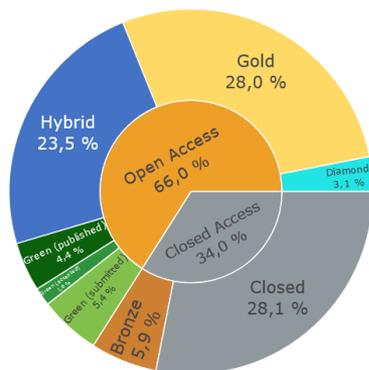
„Data is King“ - Strategische Überlegungen zum Sammelauftrag der Bibliothek im digitalen Zeitalter

Thomas Stäcker (ULB Darmstadt/FH Potsdam)

Hamburg 4.6.2024, BiblioCon

Open Access – eine Erfolgsgeschichte der Digitalisierung

Verteilung der Zeitschriftenartikel in Deutschland



Die Grafik zeigt das Open/Closed Access Verhältnis der Zeitschriftenartikel (768.540) in Deutschland für die letzten fünf Jahre (2020 - 2024) auf Basis von OpenAlex, Unpaywall sowie den im OAM verwendeten Zeitschriftenlisten (DOAJ, DOAG) an.

2/3 der Publikationen erscheinen OA

Entleihungen physischer Medien haben sich seit 2017 halbiert

Deutsche Bibliotheksstatistik (DBS) - Variable Auswertung

Trefferanzahl: 5

NR	Anzahl der ber. Bibs	Bibliothekstyp	Nr.:167 Entleihungen						
			2017	2018	2019	2020	2021	2022	2023
1	22	Regionalbibliothek	7.576.292	7.400.803	7.136.167	5.908.267	5.768.034	5.889.889	5.609.089
2	2	Nationalbibliothek	1.545.562	1.433.206	1.346.510	1.012.802	958.439	1.260.495	1.408.584
3	64	Universitätsbibliothek	42.575.406	39.976.775	37.250.590	30.244.203	26.655.320	23.271.459	20.376.968
4	102	Fach-/Hochschulbibliothek	11.965.323	11.328.803	10.717.037	7.774.679	6.529.332	5.879.977	5.478.647
5	2	Zentrale Fachbibliothek	1.558.963	1.427.407	1.282.977	1.041.392	1.011.359	786.708	660.795
Summe			65.221.546	61.566.994	57.733.281	45.981.343	40.922.484	37.088.528	33.534.083
Mittelwert			13.044.309,2	12.313.398,8	11.546.656,2	9.196.268,6	8.184.496,8	7.417.705,6	6.706.816,6

Länderbezug: Deutschland

Fragebogen: Wissenschaftliche Universal- und Hochschulbibliotheken

Quelle: OA Monitor

Die Sammlungssituation in nicht allzu ferner Zukunft

- Weitgehend alle aktuellen wissenschaftlichen Texte stehen online OA zur Verfügung und bedürfen keiner vermittelnden Bibliothek mehr
- Urheberrechtsfreie ältere Werke stehen weitgehend vollständig oder auf Abruf online OA zur Verfügung
- Der Erwerbungsetat wird überwiegend für Publikationen, nicht mehr für Käufe oder Lizenzen aufgewandt
- Archiviert werden Werke aus begrenzten Räumen im Rahmen der gesetzlichen regionalen oder nationalen Pflichtablieferung

Was bleibt für die lokale Sammlung?

- Urheberrechtsbewehrte Werke
- Werke in nicht-digitaler Form
- Pflichtexemplare
- Originale der kulturellen Überlieferung
- ...
 - Reduktion auf Archivfunktion
 - Musealisierung
 - Schwindende wissenschaftliche Relevanz

Sammlung „just in time“

„... die überwiegende Zahl der wissenschaftlichen Bibliotheken [wird]... ihren eigenen Bestand im Sinne einer ganz auf Service abstellenden Gebrauchssammlung, eine ›online core collection‹ für das Tagesgeschäft, ausrichten und letztlich nur noch ›just in time‹ »sammelnd« aktiv werden.“

(Kempf 2022)

Ende des „Gutenberg terror“

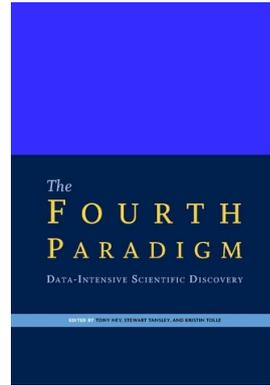
“I predict that within ten years, the days of wondering whether or not one’s library can give access to a desired resource will be a fading memory, in much the same way that reliance on a printed card catalog is a faint and fading memory in 2011.”
(Anderson 2011)

Aufstieg der Data Science

LITERATUR RECHNEN

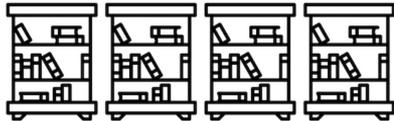
Lektüre im Computerzeitalter

(Lauer 2009)



(Hey 2009)

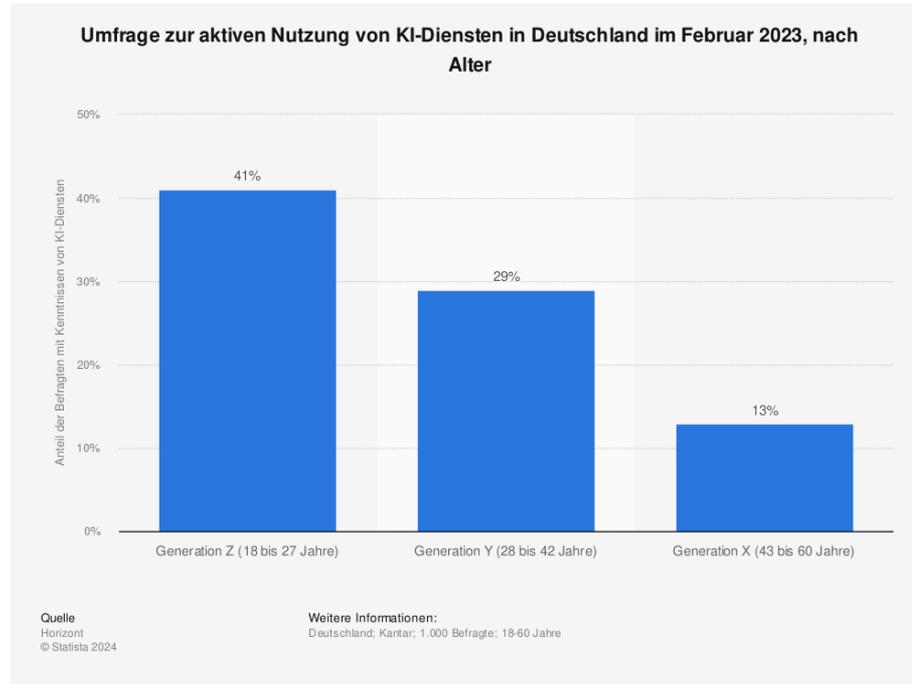
Always Already Computational: Collections as Data



(Santa Barbara Statement 2018)



Digitale Methoden im Alltag



Quelle: Statista

Data Science braucht Daten, aber ...

- in den letzten zwanzig Jahren haben Bibliotheken digitale Medien nicht gesammelt, sondern nur lizenziert (Nutzungstyp Lesen)
- Kopien von National- und Allianzlizenzen werden berechtigten Einrichtungen nur begrenzt zur Verfügung gestellt
- Archivkopien von lizenzierten digitalen Medien sind für die Wissenschaft weitgehend unzugänglich (NatHosting, u.ä.)



Quelle: Pixabay

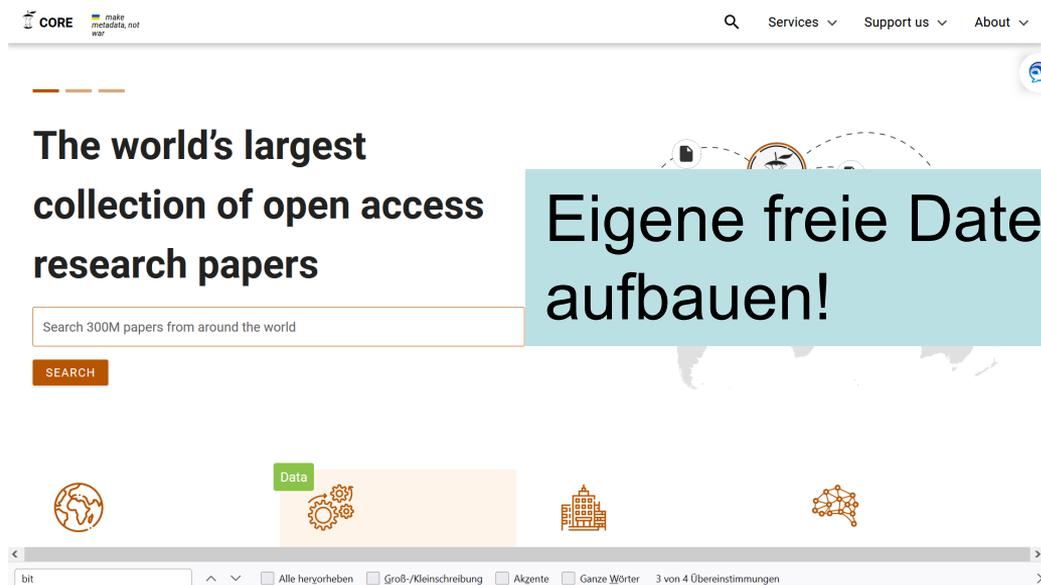
Das Ende der Sammlung?

- „Just in time“ braucht die Aggregation „just in case“
- Sammeln ist Ordnen innerhalb eines Ordnungsrahmens. Welches ist und wer schafft diese Ordnung?
- Reproduzierbarkeit und Provenienz von algorithmischen Auswertungen
- Datensouveränität



➤ FAIRe Publikationen

Kostenpflichtige Zugänge zu Sammlungen von OA Publikationen?



The screenshot shows the homepage of the CORE (Consortium of Open Repositories) website. The main heading reads "The world's largest collection of open access research papers". Below this is a search bar with the text "Search 300M papers from around the world" and a "SEARCH" button. The page features a navigation menu with "Services", "Support us", and "About". A light blue text box is overlaid on the right side of the screenshot, containing the text "Eigene freie Datensammlungen aufbauen!". At the bottom of the screenshot, there are several icons representing different data sources and a search filter bar with options like "Alle hervorheben", "Groß-/Kleinschreibung", "Akzente", "Ganze Wörter", and "3 von 4 Übereinstimmungen".

Eigene freie Datensammlungen aufbauen!

Zehn Leitgedanken zum Aufbau digitaler Sammlungen

1. Jede erworbene (lizenzierte oder OA) Publikation wird als Datei erworben. Retrodigitalisierung ist ein Erwerbungsfall.
2. Jede Publikation wird in Originalform gespeichert und mindestens zusätzlich in ein standardisiertes Basisformat konvertiert
3. Erworbene Dateien werden in der Regel dauerhaft nach OAIS Prinzipien archiviert
4. Alle relevanten Metadaten werden aus den Daten extrahiert und (semi-)automatisch standardisiert
5. Standards sind webtaugliche Standards und berücksichtigen die Maschinellebarkeit der Publikationen
6. Publikationen werden
7. Publikationen lassen sich konfigurierbaren Datensets zusammenstellen und abrufen.
8. Datensets von OA Publikationen sind auch OA
9. Datensets können zum Zwecke der Reproduzierbarkeit persistent adressiert werden
10. Digitale Sammlungen können als ganze nachgenutzt werden.

➤ **Umbau aller sammlungsbezogenen bibliothekarischen Geschäftsgänge**

Prinzipien der bibliothekarischen Zusammenarbeit

1. Alle haben alles. Keine Unikalität mehr von Beständen
2. Kooperative Aufbereitung der Publikationen und komplettes Datenset aller wissenschaftlichen Publikationen in einem zentralen Verbund oder Verbänden nach dem Modell der Verbundkataloge.
3. (Lokale) Nutzung nach dem Modell der Titelkopie.

Keine Angst vor Dubletten

„But fundamentally, you can stick a dokument in any of these archives and it will get replicated to all the other archives“

(Hey et al 2009)

- Herstellen von FAIRen Publikationen
- Aggregieren und Vereinheitlichen (*data cleaning*) von (externen) Sammlungen
- Selektion von Teilen einer Sammlung
- Transformieren in neue Derivate („KI ready“)
- Extraktion von Wissen / Informationen aus Sammlungen
- ...



»Die meisten *Bibliotheken sind entstanden*, nur wenige sind **angelegt** worden, deren sich ein so kundiger Fürst, als August war, in einer **ununterbrochenen Folge** von nahe fünfzig Jahren beeiferte.“
(Lessing 1773)

**Vielen Dank für Ihre
<emph>Aufmerksamkeit!</emph>**

