

Gefördert durch

Pestalozzianum
Stiftung für Bildung, Schule und Dialog ■■■

Digitalisierung historischer Bestände. Daten, Praktiken und Institutionen

Philipp Dreesen (Winterthur)

Peter Pohl (Innsbruck)

Von der Projekt- zur Metaperspektive

Metaperspektive auf die Diskussionen zu analogen und digitalen Wissenszugängen im explorativen interdisziplinären Projekt zur Schweizerischen Bildungsexpansion 1830–1860 (ZHAW und Universität Innsbruck, gefördert von der Stiftung Pestalozzianum, 2021–2022)

Vortrag 1: Zugang zu sozialem Wissen. Digitalisierung als Produktion und Steuerung von Gesellschaft im Bildungsdiskurs des 19. Jahrhunderts [diskurs-/computerlinguistisch]

Vortrag 2: (Re-)Kontextualisierungspraktiken in historischen Lesebüchern [literatur-/ kulturwissenschaftlich]

- > Methodologische Fragen: u.a. Korpuszusammenstellung, Repräsentativität, Kanon, ‚blinde Flecken‘
- > Fragen zu nachhaltiger Nutzung von Sammlungen und Digitalisaten

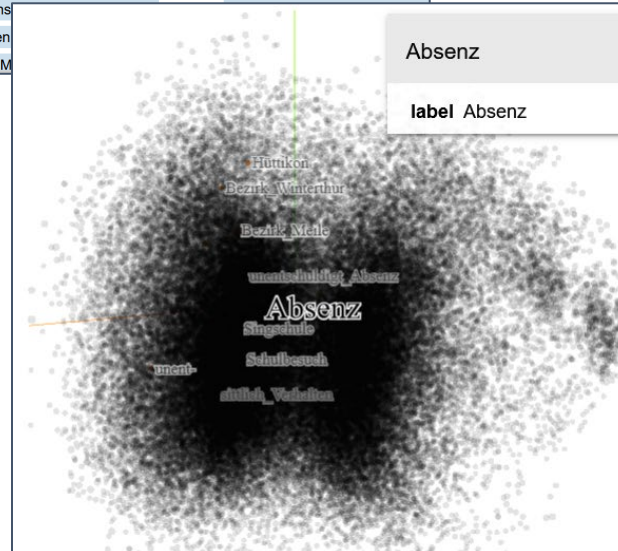
Von der Projekt- zur Metaperspektive

Projektdateien als Teil von Korpusfamilie Swiss-AL

Your query "[lemma = "Schule"]", restricted to texts meeting criteria "Structure ~text_collection~; eperiodica", returned 35,346 matches in 5,069 different texts (in 34,863,110 words [12,989 texts]; frequency: 1,013.85 instances per million words), ordered randomly [0.092 seconds]

<< >> >| Show Page: 1 Line View Show in corpus order Choose action... Go!

Text	Solution 1 to 50	Page 1 / 707
bso_002_1870_003_177_182_e47	im Kanton in 34 Volksschulen 1704 Kinder unterrichtet . Der Zustand der	Schulen wird im Allgemeinen als befriedi
apm_001_1840_016_107_111_e267	District Wald : In bessern . Zustande , als die meisten andern	Schulen des Landes finden wir schon da
slz_002_1867_012_283_290_e38	in der Gemeinde am meisten Verständniß und Herz für die Bedürfnisse der	Schule hat und auch den Lehrer kräftig
syn_001_1851_000_28_72_e108	mit immer größerem Erfolge begleitet sei . Es wird dem Urtheile mancher	Schule beigelegt , daß diese unverken
bso_002_1870_003_145_148_e40	ein Schüler hätte die neue Ortho- graphie bei seinem Austritt aus der	Schule gänzlich los ; er soll nun die Ha
ajb_001_1870_010_204_332_e377	— Walzenhausen führte , um der eingetretenen Uebervölkerung sein	Schulen in Dorf und am Platz abzuhelfer
skz_004_1863_000_309_316_e476	gestatten . Die erste Hoffnung der Or- dcnbrüder stützt sich auf diese	Schule . Nach Außen haben sie drei Ge
bso_001_1864_007_141_144_e38	Binmenstein , Oberschule , durch Lehrer Rolli „ 16. 30 Borisried ,	Schule , durch Lehrer Junker „ 8. Herzi
ass_001_1843_009_529_546_e816	gefragt wurden , ob und wie viel sie S33 freiwillig an solche	Schulen bezahlen wollten . Freiwillig bez
bsf_001_1862_002_219_225_e149	diese , namentlich für Lehrer an gemischten oder sons	
ass_001_1840_006_288_303_e707	des Mittelstandes . Auch viele Kinder von reichen	
bmb_001_1863_014_147_153_e322	Bilduna des Geistes und Herzens nur beeinträchtien M	



zhaw Digital Linguistics Workbench

🏠 Digital Discourse Lab Home

📄 Documentation

📄 CQPweb

📄 Tensorboard

Chosen Corpus
S_AL_DE_BASE_PISC_23

KWIC ▾

Corpus Query

Cooccurrences

Collocations ▾

Welcome to the Digital Linguistics Workbench!

Please get started by choosing a Corpus and Analysis method on the left sidebar menu. If your screen is too small to display the sidebar outright, please simply click or tap the menu symbol to the left of the ZHAW-Logo.

distance **COSINE** EUCLIDEAN

Nearest points in the original space:

Schulversäumnisse	0.231
unentschuldig Absenz	0.264
entschuldig	0.268
unentschuldig	0.268
Versäumnis	0.305

Vertiefte Anschlussfragen

„Distant Reading“, ein Lesen aus der Entfernung, wobei die Entfernung, das sei hier wiederholt, eine *Bedingung der Erkenntnis* ist: Sie gestattet es, Einheiten in den Blick zu nehmen, die sehr viel kleiner oder auch sehr viel größer sind als der Text: Kunstgriffe, Themen, Tropen – oder Gattungen und Systeme. Und falls einmal der Text selbst zwischen dem sehr Kleinen und dem sehr Großen verschwinden sollte, dann ist das einer der Fälle, in denen man mit Berechtigung sagen kann, dass weniger mehr ist.“

Moretti, Franco. *Distant Reading*. 2016. Konstanz: Konstanz University Press, S. 50.

Praktiken der Datenerschliessung

Grundsätzlich können auch in der Erschliessung digitaler Bestände drei Wege unterschieden werden:

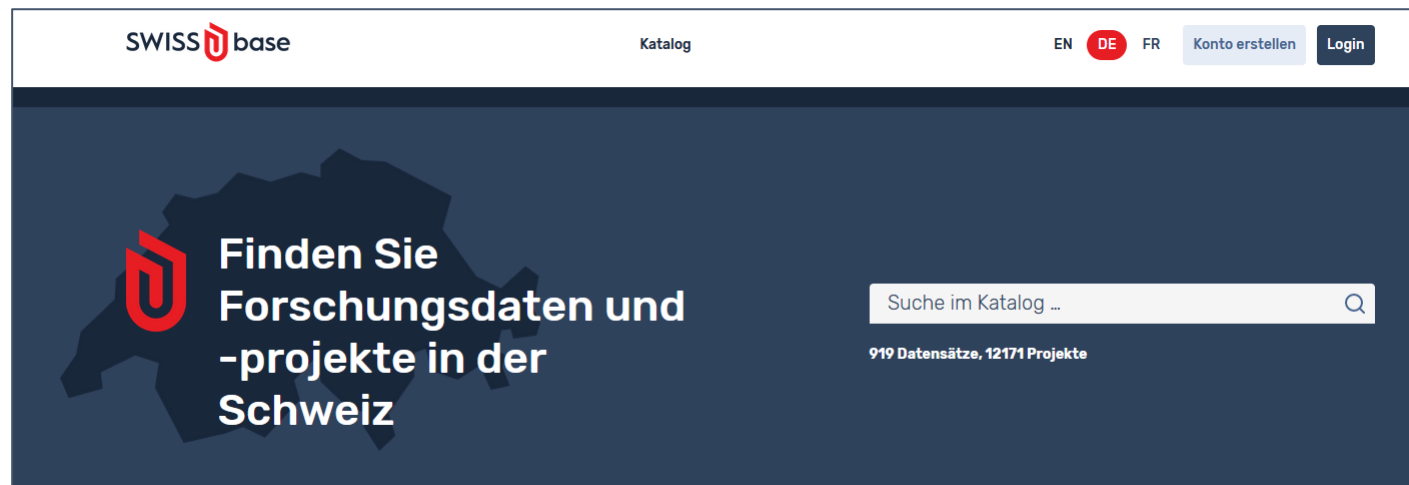
- (1) Das deduktive, hypothesenbasierte Verfahren verlangt nach standardisierten Abfragemöglichkeiten in einem Textkorpus (z.B. zur Überprüfung der relativen Nennung von Autor:innen in Lesebüchern). Learning: digitale Texterschliessung hilfreich (z.B. Wort-Frequenzen; Kollokationen zu Akteursnamen, z.B. ADJ-Attribute)
- (2) Das induktive Verfahren, das analog etwa als Grounded Theory vorkommt, wird beim Einsatz digitaler Methoden zum Auffinden von statistisch auffälligen Mustern genutzt (z.B. Topic Modelling zur Berechnung von Thematizitäten von Texten; Distribution im Zeitverlauf)
- (3) Als sinnvoll erweist sich auch in den Digital Humanities ein abduktives Vorgehen oder ein – systematischer betrachtet – trianguliertes Vorgehen, um die jeweiligen Schwächen von (1) und (2) möglichst auszugleichen. Entscheidend ist, dass auch digital erzeugte Befunde erst durch fachliche Interpretation sinnhaft und damit potenziell zu Erkenntnis werden können.

Wandel der Datennutzung



Open Science Strategien mit Open Research Data Transformation

- > Wissenstransfer von Bibliothekar:innen und Archivar:innen in die Geistes- und Kulturwissenschaften
- > Neue Stellenprofile an der Schnittstelle von Digital Humanities und Bibliothek/Archiv
- > Praktiken etwa der Standardisierung oder Katalogisierung werden zu Praktiken des Digitalen (z.B. DOI für Datensätze, Sammlungen, Forschungsergebnisse).



Praktiken des Data Life Cycle

Welche Praktiken entwickeln sich in der in der Zusammenarbeit von Digitalisierungsprojekten und Archiven, Sammlungen und Bibliotheken?

In der Forschung mit digitalen Daten zeigt sich immer wieder, dass die Fragen von findability, accessibility, interoperability und reusability (FAIR-Prinzipien) nur ansatzweise umgesetzt werden können:

- Findability: Wer ist in der Lage, die Daten zu finden? (z.B. kaum dokumentierte Daten vs. Digital Object Identifier (DOI))
- Accessibility: Wie ist ein Zugang gewährleistet? (z.B. EDU-ID vs. öffentlicher Zugang)
- Interoperability: Wie müssen Daten beschaffen sein, dass sie für Analysen bearbeitbar sind? (z.B. Scan vs. XML-Datei)
- Reusability: Wie kann eine nachhaltige Nutzung angelegt werden? (z.B. unter Verschluss halten von 'eigenen Daten' vs. frühzeitiges Teilen)

Praktiken der Verschränkung

Integration von digitalen Texten und Sammlungen

	Text	Digitalisat	Analoger Bestand
Variante 1	Text	--	--
Variante 2	annotierter (maschinell, händisch) Text	--	--
Variante 3	annotierter (maschinell, händisch) Text	Link zu gescanntem Text	
Variante 4	annotierter (maschinell, händisch) Text	Link zu gescanntem Text	Verweis auf analoges Objekt im Archiv, in Sammlung etc.

(Quellen-)Kritik des Digitalen?

„Für die Untersuchung des Inhalts und des Aussagewerts einer digitalen Quelle sind trotz der erwähnten Schwierigkeiten die gleichen Fragen wie bei der Anwendung der traditionellen Methode zu stellen.“

Föhr, Pascal. 2017. Historische Quellenkritik im Digitalen Zeitalter, S. 59. <https://doi.org/10.5451/UNIBAS-006805169>.

- Wie verhalten sich klassische Quellenkritik und digitale Quellenkritik zueinander?
- In welchem Verhältnis stehen Original und Digitalisat?

Perspektiven von Wissenschaftsforschung und Wissenssoziologie auf Digitalisierung:

- Welche Annotationsschritte sind erforderlich zur Klassifizierung von Quellen?
- Born-digital-bias: Kosten der Digitalisierung führt zu einseitiger Nutzung bereits digitalisierter Bestände.
- FAIR-bias: Gefunden und beforscht wird, was digitalisiert worden sowie zugänglich und nutzbar ist.
- Digitaler Kolonialismus: Für die grossen westlichen Sprachen liegt Analysesoftware vor, für andere weitaus weniger.

Der Wert des Analogenen im Digitalen

Digitalisierte Bestände bildungshistorischer Forschung bieten Chancen im Sinne der Open Science Bewegung:

- (1) Der Wert von historischen Daten wird gesteigert, indem sie mit neuen Daten verknüpft werden (z.B. mit Metadaten).
- (2) Durch Auffindbarkeit und standardisierte Dokumentation werden historische Daten auch für Forschungsdisziplinen, interdisziplinäre Gruppen und die Lehre interessant, die bisher nicht in diesem Forschungsfeld aktiv waren (etwa soziologische, diskursanalytische Forschung zu Lesebüchern des 19. Jahrhunderts).
- (3) Die vereinfachte Zugänglichkeit der historischen Bestände ermöglicht es, dass aktuelle Debatten etwa in der Bildungspolitik mit vergangenen Entwicklungen vergleichend perspektiviert werden können (z.B. Genderaspekte, Religion im Schulunterricht, Trigger-Warnung).