

Data Literacy

Eine Einführung für Geistes- und Kulturwissenschaftler*innen

Gero Schreier | Ursula Loosli | Universitätsbibliothek Bern

v-1.0, 2023



Data Literacy Einblick

Daten – warum ist das Thema wichtig?

Daten = "elektronisch gespeicherte
Zeichen, Angaben, Informationen"
([Duden](#))

Data Literacy Einblick

Data Literacy

"Data Literacy [Datenkompetenz] ist die Fähigkeit, planvoll mit Daten umzugehen und sie im jeweiligen Kontext bewusst einsetzen und hinterfragen zu können. Dazu gehören die Kompetenzen, Daten zu erfassen, erkunden, managen, kuratieren, analysieren, visualisieren, interpretieren, kontextualisieren, beurteilen und anzuwenden.“

(Hochschulforum Digitalisierung, 2017)

Vgl. <https://www.dataliteracy.uni-jena.de/>

Data Literacy Einblick

Ablauf

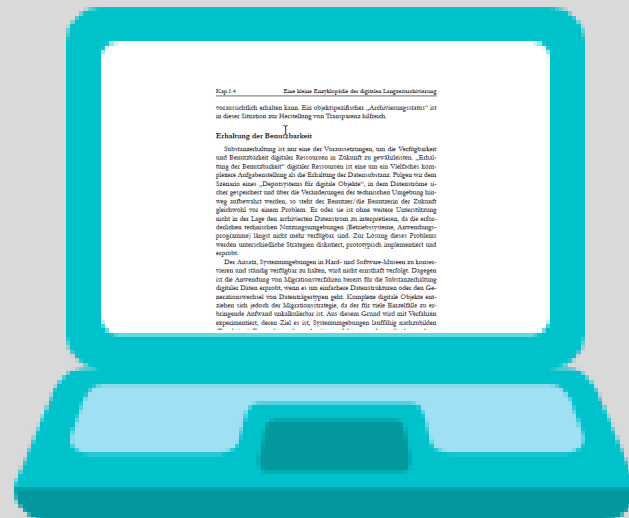
- Einführung
- Daten und beschreibende Daten (Metadaten)
- Qualitätskriterien für Datensammlungen in zwei Teilen, inkl. Übungen

Lernziele:

- Erklären können, was Daten (im geisteswissenschaftlichen) Kontext sind
- Erklären können, warum Daten standardisiert vorliegen sollten
- Die Qualität digitaler Daten beurteilen können

Data Literacy Einblick

Daten



Data Literacy Einblick

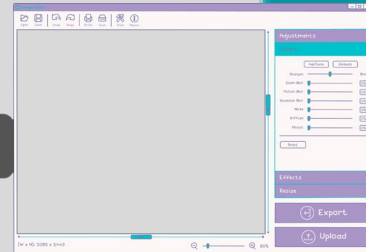
Daten als binäre Zeichen



1010101 1101110 1101001
1100010 1100101 1110010
1101110 ...

1010101 1101110 1101001
1100010 1100101 1110010
1101110 ...

u^b



Data Literacy Einblick

Beschreibende Daten

«Inhaltliche» Daten


Exp.1.4 Eine kleine Enzyklopädie der digitalen Langzeitarchivierung

voransichtlich erhalten kann. Ein objektspezifischer „Archivierungsstatus“ ist in dieser Simulation zur Herstellung von Transparenz hilfreich.

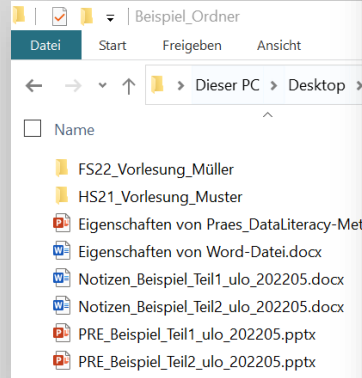
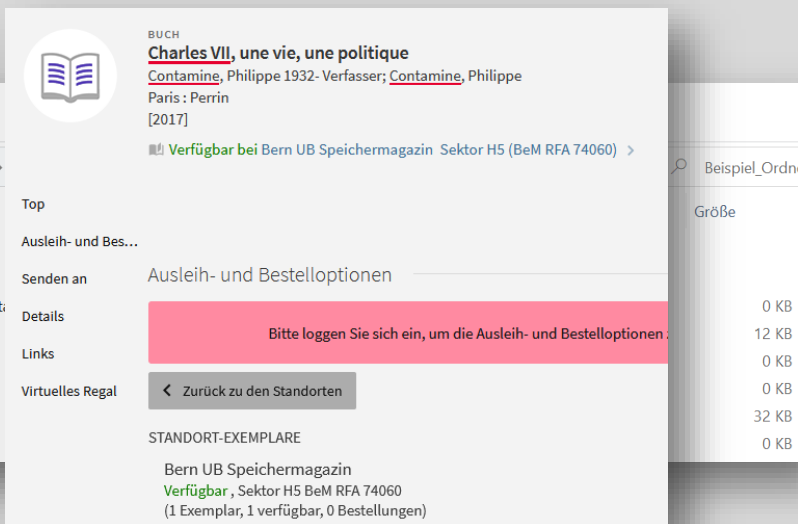
Erhaltung der Benutzbarkeit

Substanzerhaltung ist nur eine der Voraussetzungen, um die Verfügbarkeit und Benutzbarkeit digitaler Ressourcen in Zukunft zu gewährleisten. „Erhaltung der Benutzbarkeit“ digitaler Ressourcen ist eine um ein Vielfaches komplexere Aufgabenstellung als die Erhaltung der Datenstruktur. Folgen wie dem Szenario eines „Depotsystems für digitale Objekte“, in dem Datenströme sicher gespeichert und über die Veränderungen der technischen Umgebung hinweg aufrufbar werden, so teilt der Benutzer/ die Benutzerin der Zukunft gleichwohl vor einem Problem, das sich nicht durch die Erhaltung der Datenstruktur lösen lässt: Sie sind nicht in der Lage den archaischen technischen Nutzungen (z. B. durch spezielle Programme) länger nicht werden unterschiedliche Szenarien erprobt.

Der Ansatz, Systemumgebungen zu schaffen, die die Anwendung von digitalen Daten erproben, mit dem Ziel, den Benutzer/ die Benutzerin vor dem Datenverlust zu schützen, ist ein wichtiger Bestandteil der Digital Preservation. Die langfristige Erhaltung von Daten ist ein zentraler Bestandteil der Digital Preservation. Die langfristige Erhaltung von Daten ist ein zentraler Bestandteil der Digital Preservation.



Beschreibende Daten

BUCH
Charles VII, une vie, une politique
 Contamine, Philippe 1932- Verfasser; Contamine, Philippe
 Paris : Perrin
 [2017]

Verfügbar bei Bern UB Speichermagazin Sektor H5 (BeM RFA 74060) >

Top

Ausleih- und Bes...

Senden an

Details

Links

Virtuelles Regal

Ausleih- und Bestelloptionen

Bitte loggen Sie sich ein, um die Ausleih- und Bestelloptionen...

< Zurück zu den Standorten

STANDORT-EXEMPLARE

Bern UB Speichermagazin
 Verfügbar , Sektor H5 BeM RFA 74060
 (1 Exemplar, 1 verfügbar, 0 Bestellungen)

Beispiel_Ordner	Größe
	0 KB
	12 KB
	0 KB
	0 KB
	32 KB
	0 KB

Data Literacy Einblick

Beschreibende Daten (standardisiert)

The screenshot shows the Zotero application window. The left sidebar displays a hierarchical library structure under 'Meine Bibliothek'. The main pane shows a table of entries with columns for 'Titel', 'Ersteller', 'Jahr', and 'Ort'. The entry 'Charles VII, une vie, une politique' by 'Contamine' from '2017' in 'Paris' is selected. A context menu is open over this entry, listing various actions such as 'Online anzeigen', 'Schnappschuss anzeigen', 'Datei anzeigen', 'In Bibliothek anzeigen', 'Notiz hinzufügen', 'Anhang hinzufügen', 'Duplizieren', 'Eintrag aus Sammlung entfernen...', 'In Papierkorb verschieben', 'Exportieren...', 'Literaturverzeichnis erstellen' (highlighted), 'Bericht erstellen...', and 'Swissbib BB Standortabfrage'. Below the table, the 'Infos' tab is active, displaying metadata for the selected entry: 'Eintragsart: Buch', 'Titel: Charles VII, une vie, une politique', 'Autor: Contamine, Philippe', 'Zusammenfassung: Reihe: Biographie', 'Nummer der Reihe', 'Band', '# von Bänden', 'Auflage', 'Ort: Paris', 'Verlag: Perrin', 'Datum: 2017', 'Anzahl der Seiten', 'Sprache', and 'ISBN: 978-2-262-03975-2'.

Data Literacy Einblick

Beschreibende Daten (standardisiert)

Standard XY

Beispiel Daten nach Standard XY:

xy.creator: Emma Annette Wilson
xy.title: Digital humanities for librarians

Lit.-Verw.-
Programm

Anfrage (Standard XY)

Antwort (Standard XY)

Bibliothekskatalog

The screenshot shows a library catalog interface. On the left is a sidebar with a tree view of folders like 'Meine Bibliothek', 'Data Literacy', and 'Digital Humanities usw.'. The main area displays search results for the book 'Digital humanities for librarians' by Emma Annette Wilson. A detailed view is open, showing the title, author, and other metadata. The interface includes tabs for 'Infos', 'Notizen', 'Tags', and 'Zugehörig'.

The screenshot shows a detailed view of a book in a library catalog. It includes a book cover for 'Digital humanities for librarians' by Emma Annette Wilson. The text below the cover reads: 'BUCH Digital humanities for librarians Wilson, Emma 1967- (Autor_in) Lanham, MD : Rowman & Littlefield'. There is a 'Details' section with fields for 'Titel' and 'Verantwortung', both containing the book title and author name. A 'Ausgeliehen >' button is visible on the right.

Einführung Data Literacy

Datensammlungen

"Data Literacy [Datenkompetenz] ist die Fähigkeit, planvoll mit Daten umzugehen und sie im jeweiligen Kontext bewusst einsetzen und **hinterfragen** zu können.

Dazu gehören die Kompetenzen, Daten zu erfassen, erkunden, managen, kuratieren, analysieren, visualisieren, interpretieren, **kontextualisieren, beurteilen** und anzuwenden.“ (Hochschulforum Digitalisierung, 2017)

Vgl. <https://www.dataliteracy.uni-jena.de/>

Einführung Data Literacy

Datensammlungen

PROJEKT
GUTENBERG-DE

Startseite [Alle Werke](#) [Neu](#) [Information](#) [Shop](#) [Lesetips](#) [Themen](#) [Buchverlag](#) [Impressum](#) [Datenschutz](#) [Quellenangabe](#)

Autoren: [A](#) · [B](#) · [C](#) · [D](#) · [E](#) · [F](#) · [G](#) · [H](#) · [I](#) · [J](#) · [K](#) · [L](#) · [M](#) · [N](#) · [O](#) · [P](#) · [Q](#) · [R](#) · [S](#) · [T](#) · [U](#) · [V](#) · [W](#) · [X](#) · [Y](#) · [Z](#) · [Alle Autoren](#)

Heinrich von Kleist

Das Erdbeben in Chili

Heinrich von Kleist

Heinrich von Kleist

Das Erdbeben in Chili

In St. Jago, der Hauptstadt des Königreichs Chili, stand gerade in dem Augenblicke der großen Erderschütterung vom Jahre 1647, bei welcher viele tausend Menschen ihren Untergang fanden, ein junger, auf ein Verbrechen angeklagter Spanier, namens *Jeronimo Rugera*, an einem Pfeiler des Gefängnisses, in welches man ihn eingesperrt hatte, und wollte sich erheben. *Don Henrico Asteron*, einer der reichsten Edelleute der Stadt, hatte ihn ungefähr ein Jahr zuvor aus seinem Hause, wo er als Lehrer angestellt war, entfernt, weil er sich mit *Donna Josephe*, seiner einzigen Tochter, in einem zärtlichen Einverständnis befunden hatte. Eine geheime Bestellung, die dem alten Don, nachdem er die Tochter nachdrücklich gewarnt hatte, durch die hämische Aufmerksamkeit seines stolzen Sohnes verraten worden war, entrüstete ihn dergestalt, daß er sie in dem Karmeliterkloster unsrer lieben Frauen vom Berge daselbst unterbrachte.

Universitätsbibliothek

- Recherche
- E-Medien
- Datenbanken
- E-Journals
- E-Books
- E-News
- Fachgebiete

Zeitungen & Zeitschriften

- ✓ [Die Fackel \(Karl Kraus\)](#)

Datensammlungen & Verzeichnisse

- ✓ [Dokumentation zur Germanistischen Sprachwissenschaft](#)
- ✓ [German Drama Corpus](#)
- ✓ [Germanistik im Netz](#)
- ✓ [kulturfoerderung.ch](#)
- ✓ [litblogs.net - das litblog portal](#)

Virtuelle Bibliotheken

- ✓ [Deutsches Textarchiv](#)
- ! [Eighteenth Century Collections Online ECCO](#)
- ✓ [German Drama Corpus](#)

https://www.ub.unibe.ch/recherche/e_medien/datenbanken/index_ger.html

Einführung Data Literacy

Datensammlungen – Qualitätskriterien

Woran erkennen Sie,
ob ein Volltextangebot (oder eine andere
Datensammlung)
im Internet vertrauenswürdig und
für wissenschaftliche Zwecke geeignet ist?



Einführung Data Literacy

Datensammlungen – Qualitätskriterien

- Basisinformationen: **Was ist das**, woher kommt es?
- Gibt es einen **Link**, der auch in 10 Jahren noch funktioniert («Zitatsicherheit»)?
- Weiss man **eindeutig**, welche Personen, Orte und andere «Entitäten» vorkommen / erfasst sind?
- Sind die Daten leicht **zugänglich** und können sie leicht (nach-)genutzt werden?
- Werden die Daten **nach anerkannten Standards aufbereitet**, die eine flexible Nachnutzung erlauben?

Einführung Data Literacy

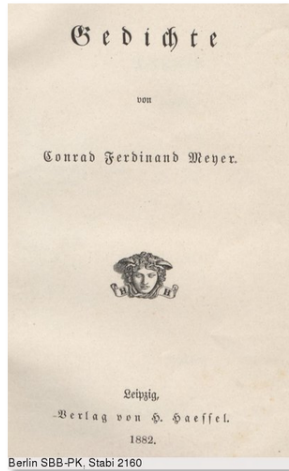
Basisinformationen I

- **Was ist das**, woher kommt es?
- Klare Basisinformation über
 - Verantwortlichkeit (Autor:in, Herausgebende, Institution)
 - Inhalt, Thema
 - zeitliche Einordnung
 - Provenienz (Herkunft)

→ Metadaten

DTA Hilfe Texte

in den Titeldaten im Korpus in der Dokumentation



Gedichte
von
Conrad Ferdinand Meyer.

Leipzig,
Verlag von F. Haessel,
1882.

Berlin SBB-PK, Stabi 2160

Text-Bild-Ansicht öffnen ...

Meyer, Conrad Ferdinand: Gedichte.
Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN:	urn:nbn:de:kobv:b4-200905193933
Titel:	Gedichte
Autor/in:	Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)
Erscheinungsjahr:	1882
Verlag/Drucker:	Haessel
Ort:	Leipzig
Auflage:	1. Auflage
Bibliothek:	Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
Signatur:	SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp:	Monographie
Verfügbarkeit:	Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0 Weitere Informationen: Nutzungsbedingungen.
Schriftart:	Fraktur
Genre:	Belletristik :: Lyrik
im DTA seit:	2008-02-26 14:31:28
Korpus:	DTA-Kernkorpus

GRUNDLAGE DIESES DIGITALISATS

Dieses Werk wurde von OCR-Software automatisch erfasst und anschließend gemäß den DTA-Transkriptionsrichtlinien von Muttersprachlern nachkontrolliert. Es wurde gemäß dem DTA-Basisformat in XML/TEI P5 kodiert.

Einführung Data Literacy

Basisinformationen II


Vertiefte Anforderungen

Metadaten sollten

- standardgebunden sein (TEI XML, Dublin Core, o.a.)
- nicht nur Basisinformationen enthalten
- relevante Links enthalten

Meyer, Conrad Ferdinand: Gedichte.
Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN: urn:nbn:de:kobv:b4-200905193933
 Titel: Gedichte
 Autor/in: Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)
 Erscheinungsjahr: 1882
 Verlag/Drucker: Haessel
 Ort: Leipzig
 Auflage: 1. Auflage
 Bibliothek: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
 Signatur:  SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp: Monographie
 Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung); CC BY-SA 4.0
 Weitere Informationen: Nutzungsbedingungen.
 Schriftart: Fraktur
 Genre: Belletristik :: Lyrik
 im DTA seit: 2008-02-26 14:31:28
 Korpus: DTA-Kernkorpus

GRUNDLAGE DIESES DIGITALISATS

Dieses Werk wurde von OCR-Software automatisch erfasst und anschließend gemäß den DTA-Transkriptionsrichtlinien von Muttersprachlern nachkontrolliert. Es wurde gemäß dem DTA-Basisformat in XML/TEI P5 kodiert.

INHALTSVERZEICHNIS

Suche im Werk

Ansichten für dieses Werk

- Text-Bild-Ansicht
- alle Faksimiles
- DTAQ (Qualitätssicherung)

Download

XML (TEI P5) · HTML · Text
 TCF (text annotation layer)
 TCF (tokenisiert, serialisiert, lemmatisiert, normalisiert)
 XML (TEI P5 inkl. att.linguistic)

Metadaten

TEI-Header · CMDI · Dublin Core

Statistiken

Scans: 353
 Zeichen: ca. 235 290
 Tokens: ca. 35 243
 Oberflächentypes: ca. 9 937

Wortwolken

- Lemmata
- Lemmata (nur Nomen)
- Types
- Types (nur Nomen)

Voyant Tools?

- transliterierter Text
- normalisierter Text
- lemmatisierter Text

Language Resource Switchboard?

Einführung Data Literacy

Zitatsicherheit I

- Gibt es einen **Link**, der auch in 10 Jahren noch funktioniert?
- Dauerhafter Identifikator, «persistent identifier» (PID)
 - DOI (Digital Object Identifier)
 - Handle
 - URN (Uniform Resource Name)



The screenshot shows the DTA (Digital Text Archive) interface. At the top, there is a search bar with the text 'suchen' and 'Hilfe'. Below the search bar, there are radio buttons for 'in den Titeldaten', 'im Korpus' (selected), and 'in der Dokumentation'. The main content area displays the title 'Meyer, Conrad Ferdinand: Gedichte. Leipzig, 1882.' and a list of bibliographic data. The URN is circled in red: 'urn:nbn:de:kobv:b4-200905193933'. The title is 'Gedichte'. The author is 'Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)'. The publication year is '1882'. The publisher is 'Haessel' in 'Leipzig'. It is the '1. Auflage' (1st edition) and is held by the 'Staatsbibliothek zu Berlin - Preussischer Kulturbesitz'. The signature is 'SBB-PK, Stabi 2160'. Below the bibliographic data, there is a section 'INFORMATIONEN ZUM WERK' with details on publication type, availability, and font.

DTA [Hilfe](#) Texte ▼

in den Titeldaten im Korpus in der Dokumentation

Meyer, Conrad Ferdinand: Gedichte.
Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN: [urn:nbn:de:kobv:b4-200905193933](https://nbn-resolving.org/urn:nbn:de:kobv:b4-200905193933)
Titel: Gedichte
Autor/in: Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)
Erscheinungsjahr: 1882
Verlag/Drucker: Haessel
Ort: Leipzig
Auflage: 1. Auflage
Bibliothek: Staatsbibliothek zu Berlin – Preussischer Kulturbesitz
Signatur: SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp: Monographie
Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung); CC BY-SA 4.0
Weitere Informationen: Nutzungsbedingungen.

Schriftart: Fraktur
Genre: Belletristik :: Lyrik
im DTA seit: 2008-02-26 14:31:28
Korpus: DTA-Kernkorpus

Einführung Data Literacy

Zitatsicherheit II

- **Persistente Identifikatoren (PID)** sind Verweise auf einen Datenbankeintrag mit der «eigentlichen» Web-Adresse (URL)
- Die URL kann sich ändern, aber der PID bleibt gleich
- URL (Uniform Resource Locator) ≠ URN (Uniform Resource Name)



DTA Hilfe

in den Titeldaten im Korpus in der Dokumentation

 Meyer, Conrad Ferdinand: Gedichte. Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN: urn:nbn:de:kobv:b4-200905193933
Titel: Gedichte
Autor/in: Conrad Ferdinand Meyer (GND, Wikipedia, ...)

Information über die URN

URN	urn:nbn:de:kobv:b4-200905193933
Verantwortliche Institution	Bibliothek der Berlin-Brandenburgischen Akademie der Wissenschaften
Erstellt / Geändert	2013-08-20 / 2021-03-18

Zugriff auf die Ressource

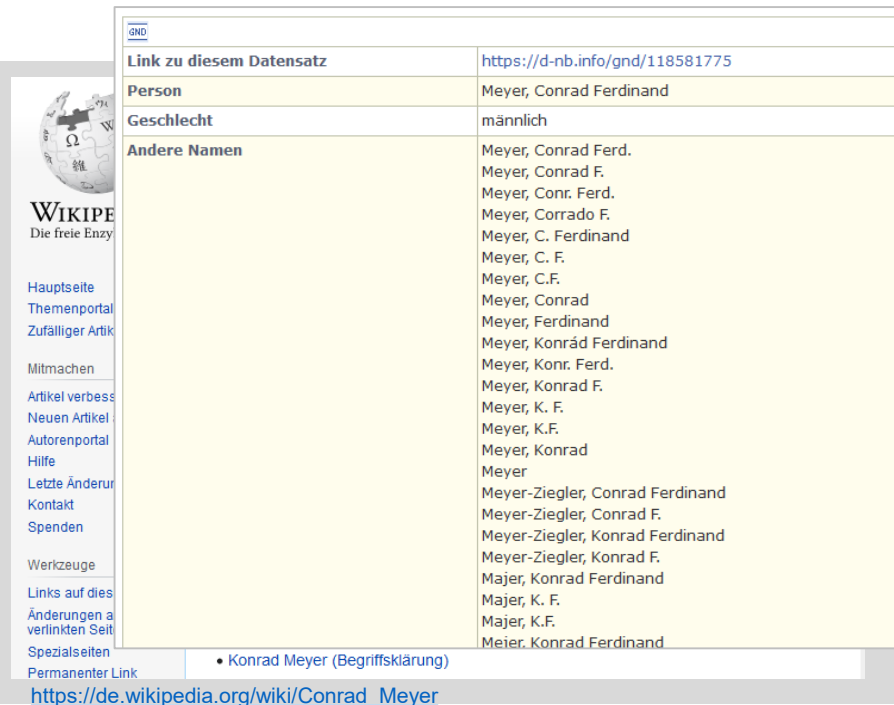
Bitte kontaktieren Sie die verantwortliche Institution, falls eine der gelisteten URLs nicht funktionieren sollte

1. URL	https://www.deutschestextarchiv.de/meyer_gedichte_1882
---------------	---

Einführung Data Literacy

Eindeutige Identifizierung I

- Weiss man **eindeutig**, welche Personen, Orte und andere «Entitäten» vorkommen / erfasst sind?
 - Vieldeutigkeit / Ambiguität vieler Namen und Begriffe
- Eindeutige Kennungen («Identifikatoren») für Personen, Orte, Institutionen
- oft Nummern, da diese eindeutiger sind



Link zu diesem Datensatz	https://d-nb.info/gnd/118581775
Person	Meyer, Conrad Ferdinand
Geschlecht	männlich
Andere Namen	Meyer, Conrad Ferd. Meyer, Conrad F. Meyer, Conr. Ferd. Meyer, Corrado F. Meyer, C. Ferdinand Meyer, C. F. Meyer, C.F. Meyer, Conrad Meyer, Ferdinand Meyer, Konrad Ferdinand Meyer, Konr. Ferd. Meyer, Konrad F. Meyer, Konrad F. Meyer, K. F. Meyer, K.F. Meyer, Konrad Meyer Meyer-Ziegler, Conrad Ferdinand Meyer-Ziegler, Conrad F. Meyer-Ziegler, Konrad Ferdinand Meyer-Ziegler, Konrad F. Majer, Konrad Ferdinand Majer, K. F. Majer, K.F. Meier, Konrad Ferdinand

WIKIPEDIA Die freie Enzyklopädie

Hauptseite
Themenportal
Zufälliger Artikel

Mitmachen

Artikel verbessern
Neuen Artikel
Autorenportal
Hilfe

Letzte Änderungen
Kontakt
Spenden

Werkzeuge

Links auf dieser Seite
Änderungen an verlinkten Seiten
Spezialseiten

Permanenter Link

• Konrad Meyer (Begriffsklärung)

https://de.wikipedia.org/wiki/Konrad_Meyer

<https://d-nb.info/gnd/118581775>

Einführung Data Literacy

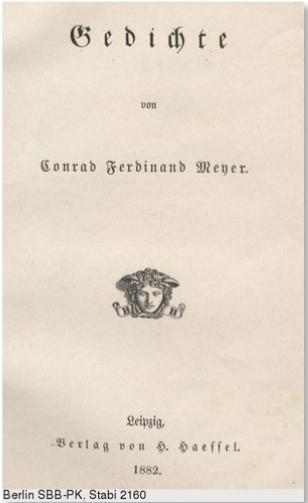
Eindeutige Identifizierung II

- Eindeutige Identifikatoren vereinfachen die Referenzierung im Internet
- **Normdaten**
 - Festlegen einer Namens-«Normalform»
 - Vernüpfung mit weiteren Informationen
- Beispiele
 - [GND: Gemeinsame Normdatei](https://d-nb.info/gnd/)
 - [Getty Geonames](https://www.getty.com/research/geonames)

GND	
Link zu diesem Datensatz	https://d-nb.info/gnd/118581775
Person	Meyer, Conrad Ferdinand
Geschlecht	männlich
Quelle	M; B 2006; B 1966; LCAuth, Wikipedia M Brümm.: Lex. dt. Dicht. IV 458 Lebensdaten: 1825-1898
Zeit	Lebensdaten: 1825-1898

D T A suchen Hilfe Texte ▾

in den Titeldaten im Korpus in der Dokumentation



Meyer, Conrad Ferdinand: Gedichte.
Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN: urn:nbn:de:kobv:b4-200905193933
 Titel: **Gedichte**
 Autor/in: **Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)**
 Erscheinungsjahr: 1882
 Verlag/Drucker: Haessel
 Ort: Leipzig
 Auflage: 1. Auflage
 Bibliothek: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
 Signatur: SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp: Monographie
 Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0
 Weitere Informationen: Nutzungsbedingungen.

Berlin SBB-PK, Stabi 2160

Einführung Data Literacy

Aufgabe 1

1) Öffnen Sie diese Seiten in Ihrem Browser:

- <https://www.projekt-gutenberg.org>
- <https://www.deutschestextarchiv.de/>
- <https://www.kleist-digital.de>

2) Suchen Sie dort nach Kleists Drama «Der zerbrochene Krug»

3) Prüfen Sie die Textangebote nach den Kriterien 1-3. Welches schneidet dabei besser ab?

Kriterien:

- Basisinformationen (Autor, Titel, zeitliche Verortung, Provenienz des Textes)
- Dauerhafter Identifikator (DOI, URN, Handle)
- Eindeutige Identifizierung von Entitäten (Personen, Orten)

Einführung Data Literacy

Zugänglichkeit

- Sind die Daten leicht **zugänglich** und können sie leicht (nach-)genutzt werden?
- Klare Darstellung auf Webseite (Usability)
- Einfache Möglichkeit zur Ansicht und zum Download
- Für Computer: Schnittstelle (API)

> [Demo-Link](#)

Meyer, Conrad Ferdinand: Gedichte. Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN:	urn:nbn:de:kobv:b4-200905193933
Titel:	Gedichte
Autor/in:	Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)
Erscheinungsjahr:	1882
Verlag/Drucker:	Haessel
Ort:	Leipzig
Auflage:	1. Auflage
Bibliothek:	Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
Signatur:	SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp:	Monographie
Verfügbarkeit:	Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0 Weitere Informationen: Nutzungsbedingungen.
Schriftart:	Fraktur
Genre:	Belletristik :: Lyrik
im DTA seit:	2008-02-26 14:31:28
Korpus:	DTA-Kernkorpus

Suche im Werk

 [Hilfe](#)

Ansichten für dieses Werk

- Text-Bild-Ansicht
- alle Faksimiles
- DTAQ (Qualitätssicherung)

Download

XML (TEI P5) · HTML · Text
TCF (text annotation layer)
TCF (tokenisiert, serialisiert, lemmatisiert, normalisiert)
XML (TEI P5 inkl. att.linguistic)

Metadaten

TEI-Header · CMDI · Dublin Core

Statistiken

Scans: 353
Zeichen: ca. 235 290
Tokens: ca. 35 243
Oberflächentypes: ca. 9 937

Einführung Data Literacy

Zugänglichkeit

- Was darf man mit den Texten eigentlich alles machen?
- So wenige rechtliche Beschränkungen wie möglich → «[Open Access](#)» statt «alle Rechte vorbehalten»

Meyer, Conrad Ferdinand: Gedichte. Leipzig, 1882.

BIBLIOGRAPHISCHE ANGABEN

URN: urn:nbn:de:kobv:b4-200905193933
Titel: Gedichte
Autor/in: Conrad Ferdinand Meyer (GND, Wikipedia, ADB/NDB)
Erscheinungsjahr: 1882
Verlag/Drucker: Haessel
Ort: Leipzig
Auflage: 1. Auflage
Bibliothek: Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
Signatur: SBB-PK, Stabi 2160

INFORMATIONEN ZUM WERK

Publikationstyp: Monographie
Verfügbarkeit: Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0
Weitere Informationen: Nutzungsbedingungen.
Schriftart: Fraktur
Genre: Belletristik :: Lyrik
im DTA seit: 2008-02-26 14:31:28
Korpus: DTA-Kernkorpus

Suche im Werk

Hilfe

Ansichten für dieses Werk

- Text-Bild-Ansicht
- alle Faksimiles
- DTAQ (Qualitätssicherung)

Download

XML (TEI P5) · HTML · Text
TCF (text annotation layer)
TCF (tokenisiert, serialisiert, lemmatisiert, normalisiert)
XML (TEI P5 inkl. att.linguistic)

Metadaten

TEI-Header · CMDI · Dublin Core

Statistiken

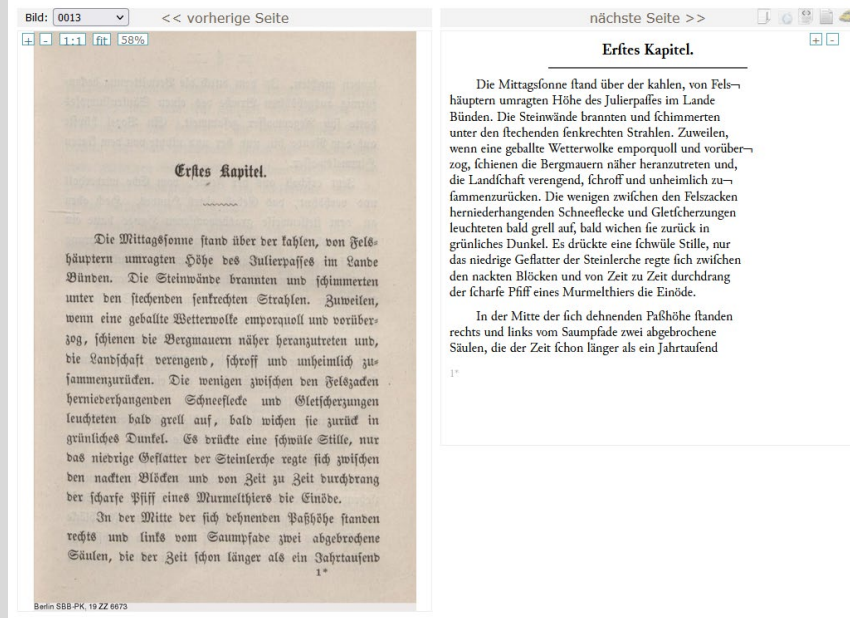
Scans: 353
Zeichen: ca. 235290
Tokens: ca. 35243
Oberflächentypes: ca. 9937

Einführung Data Literacy Standards

- Werden die Daten **nach anerkannten Standards aufbereitet**, die eine flexible Nachnutzung erlauben?
- Für Text-Daten: TEI-XML
 - formatunabhängige Einbettung von Informationen (inhaltlich, Layout) → «Annotationen»
 - kann in verschiedenen Formaten (PDF; HTML) dargestellt werden

```
</head>  
<milestone unit="section" rendition="#hr"/>  
<fw type="sig" place="bottom">  
  <hi rendition="#g">Meyer</hi>  
  , Georg Jenatsch. 1  
</fw>
```

Meyer, Conrad Ferdinand: Georg Jenatsch. Leipzig, 1876.



Einführung Data Literacy

Aufgabe 2

1) Öffnen Sie diese Seiten in Ihrem Browser:

- <https://www.projekt-gutenberg.org>
- <https://www.kleist-digital.de>
- <https://www.deutschestextarchiv.de/>

2) Suchen Sie dort nach Kleists Drama "Der zerbrochene Krug"

3) Prüfen Sie die Textangebote nach den Kriterien 4-5. Welches schneidet dabei besser ab?

Kriterien:

- Basisinformationen (Autor, Titel, zeitliche Verortung, Provenienz des Textes)
- Dauerhafter Identifikator (DOI, URN, Handle)
- Eindeutige Identifizierung von Entitäten (Personen, Orten)
- Leichte Zugänglichkeit und (Nach-)Nutzbarkeit (Lizenz, Downloadmöglichkeit)
- Aufbereitung nach Standards (TEI-XML)

Einfü Ausb

The screenshot displays the Voyant Tools interface for a corpus of 1 document with 5,389 unique word forms. The interface includes a navigation bar at the top with tabs for 'Cirrus', 'Begriffe', 'Links', 'Reader', and 'Worttrauben'. The main content area is divided into several sections:

- Word Cloud:** A word cloud on the left side, with the most prominent words being 'don', 'fernando', 'josephe', 'jeronimo', 'domna', and 'menschen'.
- Text Snippet:** A central text snippet titled 'Das Erdbeben in Chili; Heinrich von Kleist: Werke und Briefe in vier Bänden...' with a sub-header 'Das Erdbeben in Chili'. The text describes the earthquake in St. Jago, Chile, in 1647, mentioning the death of a young man and the actions of Don Henrico Asteron and Don Fernando.
- Line Graph:** A line graph on the right side showing 'relative Frequenzen' (relative frequencies) across 'Dokument Segmente' (document segments). The graph features five lines representing different names: don (blue), donna (magenta), fernando (green), jeronimo (purple), and josephe (cyan).
- Concordance Table:** A table at the bottom right showing the context of the word 'don' in the document. The table has columns for 'links', 'Begriff', and 'rechts'.

Summary Statistics:

- Dieser Korpus hat 1 Dokument mit 5,389 einzigartige Wortformen. Erstellt vor 3 Sekunden.
- Wortschatzdichte: 0.332
- Readability Index: 15.048
- Durchschnittliche Wörter pro Satz: 28.2
- Die häufigsten Begriffe im Korpus: don (40), fernando (27), jeronimo (26), domna (25), josephe (24)

Concordance Table:

Dokument	links	Begriff	rechts
Das Erd...	hatte, und wollte sich erkennen.	don	Henrico Asteron, einer der reichsten
Das Erd...	geheime Bestellung, die dem alten	don	, nachdem er die Tochter nachdrücklich
Das Erd...	schwieg – aus einem andern Grunde,	don	Fernando; in diesen schrecklichen Zeiten
Das Erd...	legte ihn an ihre Brust.	don	Fernando war sehr dankbar für
Das Erd...	das innigste und zärtlichste von	don	Fernandos beiden Schwägerinnen, die sie
Das Erd...	kannte, empfangen ward. Donna Elvire,	don	Fernandos Gemahlin, welche schwer in
Das Erd...	Freundlichkeit zu sich nieder. Auch	don	Pedro, sein Schwiegervater, der an
Das Erd...	in Strömen zur Stadt. In	don	Fernandos Gesellschaft ward die Frage
Das Erd...	Messe hören sollte, und rief	don	Fernando auf, die Gesellschaft zu
Das Erd...	ihn wieder still. Hierauf bot	don	Fernando, dem die ganze Würdigkeit
Das Erd...	mit Donna Elvire gesprochen hatte;	don	Fernando! rufen hörte, und dem
Das Erd...	mit unruhigen Tritten nachhellen sah.	don	Fernando hielt, und kehrte sich

u^b

UNIVERSITÄT
Erlangen
<https://voyant-tools.org/?input=https://textgridlab.org/1.0/tgcrud-public/fest/textgrid:r0x.5.1gata>

Data Literacy Einblick

Fragen?

Data Literacy Ausblick

Weiterführende Kurse & Angebote der UniBE

- [Data Literacy Selbstlernmodul](#): Einblick in Metadaten, Datenstandards, Dateiformate und Möglichkeiten der Dateistrukturierung (UB)
- [Data Literacy für Geisteswissenschaftler:innen](#) (Digital Humanities, jeweils im FS)
- [Grundkurs Programmieren](#) (für alle Studierenden der UniBE)

Data Literacy Ausblick

Quellen und weiterführende Hinweise

- CrashCourse: Representing Numbers and Letters with Binary: Crash Course Computer Science #4, 15.03.2017. Online: <<https://www.youtube.com/watch?v=1GSjbWt0c9M>>
- CrashCourse: Files & File Systems: Crash Course Computer Science #20, 13.07.2017. Online: <https://www.youtube.com/watch?v=KN8YgJnShPM>
- Marie Flüh (2019): „Textvisualisierung mit Voyant“. In: forTEXT. Literatur digital erforschen. URL: <https://fortext.net/routinen/lerneinheiten/textvisualisierung-mit-voyant> [Zugriff: 01. Juni 2023].
- GO FAIR, FAIR Principles, <https://www.go-fair.org/fair-principles/> (31. Mai 2023)
- Henny, Ulrike; Neuber, Frederike: Criteria for Reviewing Digital Text Collections, version 1.0, 02.2017, <<https://www.i-d-e.de/publikationen/weitereschriften/criteria-text-collections-version-1-0/>> (29.03.2023).

Data Literacy Ausblick

Quellen und weiterführende Hinweise

- Neuber, Frederike; Schulz, Daniela: EDITORIAL: FAIR Enough? Evaluating Digital Scholarly Editions and the Application of the FAIR Data Principles, in: RIDE – A review journal for digital editions and resources 16, 2023. <https://doi.org/10.18716/RIDE.A.16.0> (29.03.2023).
- Sahle, Patrick: Kriterien für die Besprechung digitaler Editionen V. 1.1, Institut für Dokumentologie und Editorik, 06.2014, <https://www.i-d-e.de/publikationen/weitereschriften/kriterien-version-1-1/> (29.03.2023).
- Schöch, Christof: Big? Smart? Clean? Messy? Data in the Humanities, in: Journal of Digital Humanities 2 (3), 2013. Online: <<http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>>, Stand: 07.06.2021.
- Zeng, Marcia Lei; Qin, Jian: Metadata, 3rd ed., London 2022.

Data Literacy Einführung

Vielen Dank für Ihre Aufmerksamkeit!

Gero Schreier gero.schreier@unibe.ch

Ursula Loosli ursula.loosli@unibe.ch