

**Integration einer Sammlung
historischer, deutschsprachiger Texte
in die Infrastruktur des Deutschen Textarchivs
an der Berlin-Brandenburgischen Akademie der Wissenschaften**

Marko Neumann (Justus-Liebig-Universität Gießen)

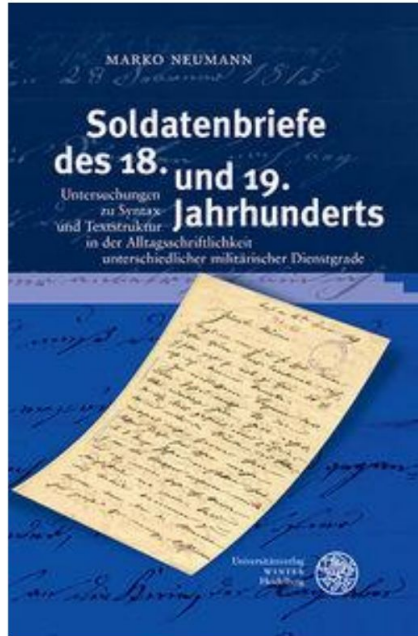
Marius Hug (Berlin-Brandenburgische Akademie der Wissenschaften)

Text+

Wohin damit? Storing and reusing my language data

Mannheim, 22.06.2023

1. Die Dissertation



Soldatenbriefkorpus

Kriterien der Zusammenstellung:

- Textsorte: Privatbrief
- Verfasser: Soldat
- Adressat: Familie
- Schreibanlass: Krieg
- Sprache: deutsch
- Überlieferung: Original

Prinzip der größtmöglichen Varianz

1. Die Dissertation

1.1 Beschaffung der Korpustexte:



<https://depositonce.tu-berlin.de/bitstreams/9ec132e6-abe7-42f6-8231-503e704f40a9/download>

Archivauswahl:

- Stadtarchiv Stralsund (3 Briefe)
- Staatsbibliothek zu Berlin Preußischer Kulturbesitz (25 Briefe)
- Museumsstiftung Post und Telekommunikation Berlin (6 Briefe)
- Geheimes Staatsarchiv Preußischer Kulturbesitz Berlin (8 Briefe)
- Stadtarchiv Braunschweig (9 Briefe)
- Niedersächsisches Landesarchiv Wolfenbüttel (39 Briefe)
- Bayerisches Hauptstaatsarchiv München – Abteilung IV: Kriegsarchiv (31 Briefe)
- Österreichisches Staatsarchiv Wien – Kriegsarchiv (23 Briefe)

1. Die Dissertation

1.2 Soldatenbriefkorpus (170 Briefe)

Zeitliche Verteilung:

1745 bis 1800: 18 Briefe

1801 bis 1850: 42 Briefe

1851 bis 1872: 110 Briefe

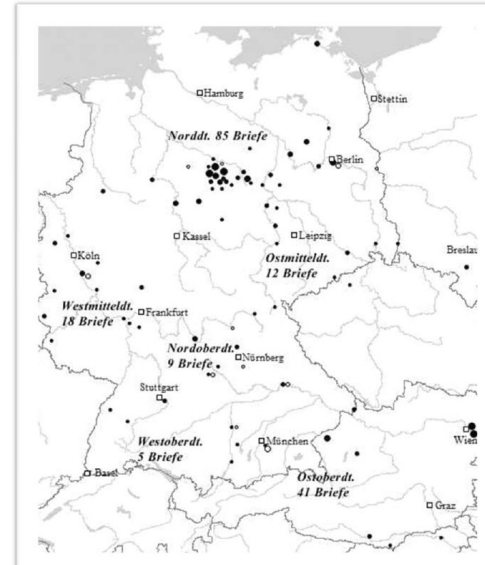
Verteilung nach Dienstgrad:

einfache Soldaten: 95 Briefe

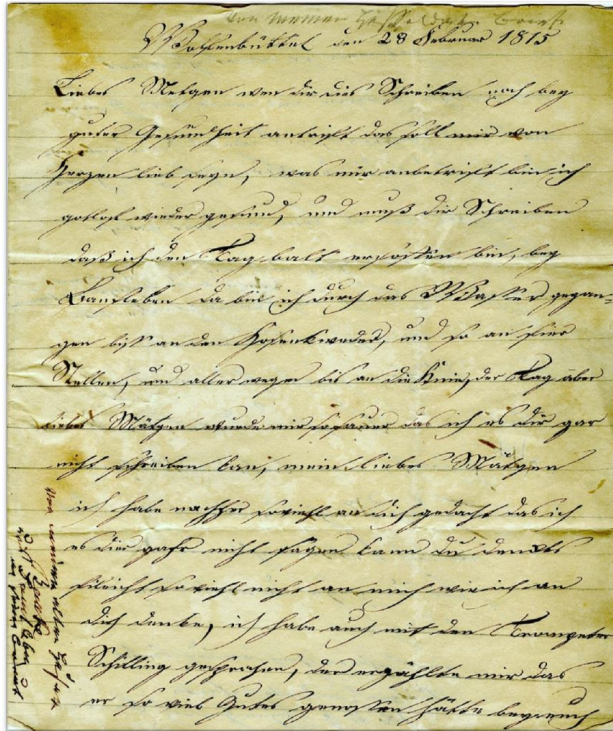
Unteroffiziere: 24 Briefe

Offiziere: 51 Briefe

Räumliche Verteilung:



1. Die Dissertation



Johann H. Baake, 28.02.1815 (Stadtarchiv Braunschweig: H VI 6:11)

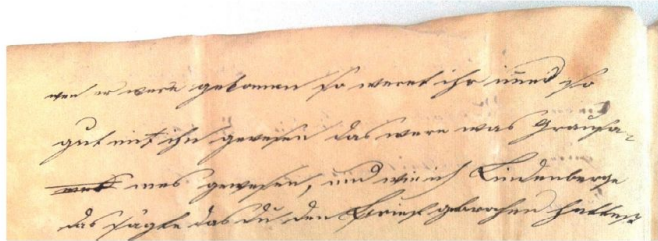
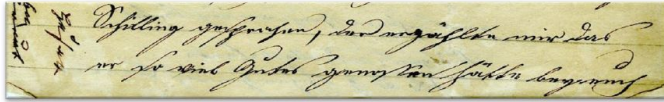
Wolfenbüttel den 28 Februar 1815

Liebes Metgen wen dir dies Schreiben noch bey guter Gesundheit antrifft das soll mir von Herzen lieb seyn, was mir anbetrifft bin ich gotlob wieder gesund, und muß dir Schreiben daß ich den Tag balt ersoffen bin, bey Bansleben da bin ich durch das Wasser gegangen biss an den Hosenkweder, und so an fier Stellen, und aller wegen bis an die Knie, der Tag aber Liebes Mätgen wurde mir so sauer das ich es dir gar nicht schreiben kan, mein liebes Mätgen ich habe nachher soviehl an dich gedacht das ich es dier gahr nicht sagen kann du dencks fileicht so viehl nicht an mich wie ich an dich denke, ich habe auch mit den Trompeter Schilling gesprochen, der erzählte mir das er so viel Gutes genossen hätte bey euch [1]

Von meinem alten Husar
Baake
aus Samtleben
im steten Beruf
[Randbemerkung 1]

1. Die Dissertation

1.3 Bestimmung von Ganzsatzgrenzen mithilfe von Layoutspezifika



Schilling gesprochen, || der erzählte mir | das
er so viel Gutes genossen hätte bey euch || (?)
[1]

wen er were gekommen | so weret ihr immer so
gut mit ihn gewesen | das were was grausam
mes mes gewesen, und wie ich Lindenberge
das sagte das du den Brief gebrochen hattest

1. Die Dissertation

1.4 Entgeltordnung für das Niedersächsische Landesarchiv

4.	Nutzungs- und Veröffentlichungsgenehmigungen (Erlaubnis zur Veröffentlichung oder Vervielfältigung von Reproduktionen von Archivgut zu gewerblichen oder geschäftlichen Zwecken. Ansprüche Dritter aus Urheber-, Verwertungs- und Lizenzrechten sind gesondert abzugelten.)	
4.1	in Büchern, Broschüren, Zeitschriften und Zeitungen je nach Art und Auflage	
	a) bis 5 000 Exemplare	40,00
	b) bis 10 000 Exemplare	100,00
	c) für jede weiteren 1 000 Exemplare	10,00
	bis zu einem Höchstbetrag von	500,00
4.2	auf Plakaten und Ansichtskarten	das Doppelte des Entgelts nach Nummer 4.1
4.3	bei Neuauflagen und Nachdrucken	die Hälfte des Entgelts nach Nummer 4.1
4.4	für die Verwendung im Film oder Fernsehen	je angefangene Minute 100,00
4.5	für die Verwendung auf Datenträgern	wie Nummer 4.1
	Bei gleichzeitiger Verwendung in gedruckten Publikationen ermäßigt sich das Entgelt für die Verwendung auf Datenträgern auf die Hälfte.	
4.6	Einblendung in Onlinedienste, Internetpräsentationen und vergleichbare Medien je Reproduktion	
	a) für bis zu einem Monat	40,00
	b) für sechs Monate	100,00
	c) für ein Jahr	150,00

1. Die Dissertation

Universitätsverlag
WINTER
Heidelberg

Anmelden AGB und Widerrufsbelehrung Datenschutz Versandkosten Impressum  

Suche nach Titel, Autor, ISBN, ... 
Erweiterte Suche

Aktuell Über uns Kontakt vor Ort Partner Rechte/Lizenzen    (0) 

- Fachgebiete
- Buchreihen
- Zeitschriften
- E-Books
- Open Access
-  Edition S
- Kataloge
- Angebote



Soldatenbriefe des 18. und 19. Jahrhunderts

Untersuchungen zu Syntax und Textstruktur in der Alltagsschriftlichkeit unterschiedlicher militärischer Dienstgrade

Marko Neumann

1. Auflage, 2019
365 Seiten

ISBN: 978-3-8253-4642-3
Sortiment: Buch
Ausgabe: Gebunden
Fachgebiet: Germanistik
Reihe: Germanistische Bibliothek, Band: 68
lieferbar: 27.09.2019

52,00 €*

 In den Warenkorb legen

Schlagwörter: *Krieg, Soziolinguistik, Textlinguistik, Syntax, 18. Jahrhundert, 19. Jahrhundert, Neuhochdeutsch, Soziogrammatik, Privatbriefe, Briefsteller, Soldatenbriefe, Alltagsschriftlichkeit, Briefkorpus, Musterbriefe, Militär*

Soldatenbriefe des 18. und 19. Jahrhunderts eröffnen einen sprachhistorisch bislang kaum erprobten Zugang zur privaten Alltagsschriftlichkeit von Offizieren, Unteroffizieren und einfachen Soldaten. Stimmen die Schreibenanlässe im Krieg auch für alle Dienstgradgruppen weitgehend überein, erweist sich die Schreiberfahrung aufgrund des Bildungshintergrunds doch als deutlich verschieden.

Die Studie arbeitet syntaktische und textuelle Gemeinsamkeiten und Unterschiede in den Briefen der Schreibergruppen heraus und lotet damit die Varianzbreite historischer Alltagsschriftlichkeit im Untersuchungszeitraum aus. Quellengrundlage ist ein umfangreiches Korpus von Soldatenbriefen aus dem gesamten deutschsprachigen Raum. Ein Vergleichskorpus mit Musterbriefen aus zeitgenössischen Briefstellern ermöglicht darüber hinaus die Kontrastierung des tatsächlichen Sprachgebrauchs mit einschlägigen präskriptiven Texten.

[Rezensionen](#) [Dokumente](#)

[Anhang zu Marko Neumann: Soldatenbriefe des 18. und 19. Jahrhunderts](#) (PDF 4152kB)

1. Die Dissertation



H / SOZ / KULT

Kommunikation und Fachinformation
für die Geschichtswissenschaften

Abonnieren /
Einreichen /

[Home](#) [Chancen](#) [Forum](#) [Forschungsberichte](#) [Miscellen](#) [Rezensionen](#) [Termine](#) [Tagungsberichte](#) [Zeitschriften](#) [Suchen](#)

Neuere Geschichte /

M. Neumann: Soldatenbriefe des 18. und 19. Jahrhunderts

[Info](#) 
[Drucken](#) 
[PDF](#) 

[◀ zurück](#) [3/20 Einträge](#) [vorwärts ▶](#)

Titel	Soldatenbriefe des 18. und 19. Jahrhunderts. Untersuchungen zu Syntax und Textstruktur in der Alltagsschriftlichkeit unterschiedlicher militärischer Dienstgrade
Autor(en)	Neumann, Marko
Reihe	Germanistische Bibliothek 68
Erschienen	Heidelberg 2019: Universitätsverlag Winter
Anzahl Seiten	IX, 365 S.
Preis	€ 52,00
ISBN	978-3-8253-4642-3

Rezensiert für H-Soz-Kult von

Anke Fischer-Kattner, Historisches Institut, Universität der Bundeswehr München

Marko Neumann schlägt mit seiner Rostocker Dissertation eine Brücke zwischen Militärgeschichte und Germanistik – angesichts der Distanz zwischen den

Redaktion

Veröffentlicht am
05.06.2020

Autor(en)
[Anke Fischer-Kattner](#)

Beiträger
[Anke Fischer-Kattner](#)

Redaktionell betreut durch
[Daniel Benedikt Stienen](#)

Zitation 

Klassifikation

2. NFDI/Text+

Text- und sprachbasierte Forschungsdaten im Verbund Text+ (<https://www.text-plus.org/>)



2. NFDI/Text+



DE

Support | Helpdesk

About Us ▾

Research Data ▾

Networking ▾

Links & Downloads

News & Events ▾



Text+ User Story

Soldiers' letters of the 18th and 19th centuries: From the PDF edition to reusable, interoperable research data

Marko Neumann (Justus-Liebig Universität Gießen), via Christian Thomas (Berlin-Brandenburgische Akademie der Wissenschaften)

DFG subject area: [104 Linguistics](#)

Text+ data domain: [Collections](#)

Motivation

For the investigation of linguistic phenomena, dedicated corpora of certain types of texts, periods of time or authorships are required. Due to the lack of available data for his research on everyday writing, in particular on the syntax in soldiers' letters of the 18th and 19th century, Dr. Marko Neumann has created such a corpus especially for his doctoral thesis. To this end he has transcribed unique sources in laborious archival work, supplemented by selected letters from older, sufficiently reliable editions. This alone is not the common practice, even less is it common that such data is then made freely available for subsequent use for other researchers—among other things because this takes a lot of effort, but at the same time brings comparatively little reputation in return.

3. Das Deutsche Textarchiv (DTA)

Kernkorpus (150 Mio. Token)

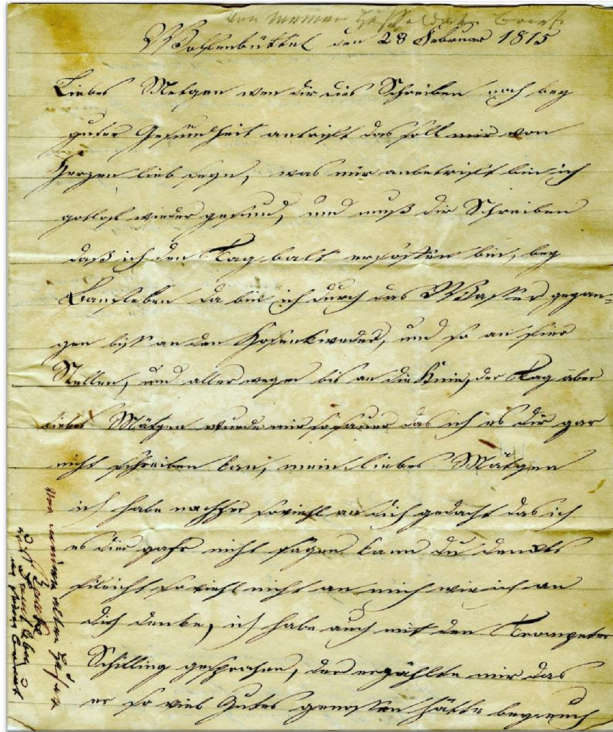
- DFG-geförderte Digitalisierung (double-keying) von rund 1.500 deutschen Werken des 17. bis 19. Jahrhunderts
- Bereitstellung als Korpus der neuhochdeutschen Sprache

DTA-Erweiterungen (220 Mio. Token)

- Aktives Archiv: Übernahme und Kuration von digitalisierten Sammlungen
- Öffnung zur Community

Weitere: Dinger (77.5 Mio.), DSDK (53.7 Mio.), DiBiLit (110 Mio.), Reichstag (311 Mio.), Brüdergemeinde (20 Mio.), Pitaval (7 Mio.), Jean Paul Briefe (1 Mio.),
Soldatenbriefe (0.1 Mio.)

3. Das Deutsche Textarchiv (DTA)



Johann H. Baake, 28.02.1815 (Stadtarchiv Braunschweig: H VI 6:11)

Wolfenbüttel den 28 Februar 1815

Liebes Metgen wen dir dies Schreiben noch bey guter Gesundheit antrifft das soll mir von Herzen lieb seyn, was mir anbetrifft bin ich gotlob wieder gesund, und muß dir Schreiben daß ich den Tag balt ersoffen bin, bey Bansleben da bin ich durch das Wasser gegangen biss an den Hosenkweder, und so an fier Stellen, und aller wegen bis an die Knie, der Tag aber Liebes Mätgen wurde mir so sauer das ich es dir gar nicht schreiben kan, mein liebes Mätgen ich habe nachher soviehl an dich gedacht das ich es dier gahr nicht sagen kann du denkts fileicht so viehl nicht an mich wie ich an dich denke, ich habe auch mit den Trompeter Schilling gesprochen, der erzählte mir das er so viel Gutes genossen hätte bey euch [1]

Von meinem alten Husar
Baake
aus Samtleben
im steten Beruf
[Randbemerkung 1]

3. Das Deutsche Textarchiv (DTA)

```
<sourceDesc>
  <bibl type="MAN">Soldatenbrief Nr. 169 von Michael Eimgartner, 1871.</bibl>
  <biblFull>
    <titleStmt>
      <title level="m" type="main">
        [SB-Nr. 169: Michael Eimgartner (Mannschaften?), 1871, Ostoberdeutsch]
      </title>
    </titleStmt>
    <author>
      <persName role="Mannschaften" cert="low">
        <surname>Eimgartner</surname>
        <forename>Michael</forename>
      </persName>
    </author>
  </titleStmt>
  <editionStmt>
    <edition/>
  </editionStmt>
  <publicationStmt>
    <publisher/>
    <date when="1871">1871</date>
  </publicationStmt>
</biblFull>
<msDesc>
  <msIdentifier>
    <repository>
      Bayerisches Hauptstaatsarchiv – Abteilung IV (Kriegsbriefe 235)
    </repository>
  </msIdentifier>
  <physDesc/>
</msDesc>
</sourceDesc>
```

<sourceDesc> des TEIHeaders von Soldatenbrief Nr. 169.

3. Das Deutsche Textarchiv (DTA)

DTABf – Das TEI-Standardformat des DTA

- TEI = De-facto-Standards zur Kodierung von gedruckten Werken innerhalb der Geisteswissenschaften
- Basisformat des Deutschen Textarchiv (DTABf) = projektspezifisches TEI-Subset zur Auszeichnung von sprachlichen Informationen

```
<rs type="propNounPlaceName">Marbach</rs>, <name type="place">Weimar</name> und  
<placeName role="venue">Wolfenbüttel</placeName> sind Orte in Deutschland.
```

```
<placeName ref="https://www.geonames.org/2873776/">Marbach</placeName>,  
<placeName ref="https://www.geonames.org/2812482/">Weimar</placeName> und  
<placeName ref="https://www.geonames.org/2806914/">Wolfenbüttel</placeName>  
sind Orte in Deutschland.
```

=> DTABf hilft dabei, Ambiguitäten zu vermeiden
Interoperabilität (mit anderen Ressourcen) zu verbessern

3. Das Deutsche Textarchiv (DTA)

The screenshot shows a GitHub repository page for 'Frank Wiegand cleanup'. At the top, there are navigation buttons for 'main', '2 branches', and '0 tags'. Below this is a table of files and folders:

File/Folder	Commit Message	Last Commit
attic	initial commit	last year
data	add msidentifier injection	last year
scripts	add msidentifier injection	last year
LICENSE.md	initial commit	last year
README.md	Update README.md	10 months ago

Below the table, the 'README.md' file is open, showing the title 'Corpus Soldatenbriefe (1745–1872)'. The text describes the corpus as 'Soldiers' Letters' comprising 170 letters from officers, non-commissioned officers, and ordinary soldiers, addressed to their families back home. It mentions the focus on letters from the Coalition and Liberation Wars (1792–1815), the German War (1866), and the Franco-Prussian War (1870/71). It also notes that the writing experience of the three rank groups (officer, non-commissioned officer, enlisted men and common soldier) is very different. The corpus thus offers the possibility of examining the entire breadth of historical everyday writing on the basis of a single type of letter. A link to 'Text+' is provided at the end of the text.

Soldatenbriefe

Marius Hug edited this page on Jun 3, 2022 · 4 revisions

- **Titel:** Soldatenbriefe des 18. und 19. Jahrhunderts
- **Umfang:** 170 Briefe
- **Sprache:** (Früh-)Neuhochdeutsch
- **Textgattung:** Gebrauchsliteratur/Brief
- **Kontakt:** Dr. Marco Neumann
- **Text+ User-Story:** <https://www.text-plus.org/en/research-data/user-story-508/>
- **Kuration/Konvertierung:**
 - BBAW/Text+
 - Christian Thomas, Frank Wiegand, Marius Hug
- **Lizenz:** CC BY-SA 4.0

Vorlage

Die Datenbasis für die Konvertierung und Integration des gesamten Soldatenbrief-Korpus bestand aus einer docx-Datei. Diese enthielt in einer einspaltigen Tabelle insgesamt 170 Briefe. Desweiteren wurde eine docx-Datei übergeben, die ebenfalls in einer Tabelle (7-spaltig) als Verzeichnis der untersuchten Texte die zugehörigen Metadaten enthielt.

Workflow

Infrastruktur

Für die kollaborative Arbeit an den Daten wurde ein GitHub-repository angelegt: <https://github.com/deutschestextarchiv/soldatenbriefe>.

Das Repository enthält neben einer **README** und **LICENSE** verschiedene Unterordner:

- **attic:** Ausgangsdateien
- **data:** die einzelnen Dateien des Korpus
- **scripts:** Skripte, Tools und Metadaten

The screenshot shows a GitHub Wiki page for 'Soldatenbriefe'. It features a search bar at the top with the text 'Find a page...'. Below the search bar, there is a list of pages under the heading 'Soldatenbriefe':

- Vorlage
- Workflow
- Infrastruktur
- Erstellung eines DTABf-konformen TEI-Headers für das gesamte Korpus
- Vorbereitung der Kuration
- Kuration
- Korpus-Bau
- Qualitätssicherung
- Wohlfelgformtheit
- Validität
- Publikation
- GitHub
- Landingpage
- Zenodo

Clone this wiki locally

<https://github.com/deutschestex>

Soldatenbriefe im git-Repository (<https://github.com/deutschestextarchiv/soldatenbriefe>) inkl. Dokumentation zur Datenaufnahme und -bereitstellung.

3. Das Deutsche Textarchiv (DTA)

Akademienvorhaben Johann Friedrich Blumenbach – online
1779–1832

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Publikations JA

📄 Inkunabeln 23 256281

📄 Inkunabeln 1276522

📄 Genre Wissenschaft CC BY-SA 4.0

Das Korpus „Blumenbach online“ enthält Publikationen und die naturhistorischen Sammlungen des Göttinger Naturforschers Johann Friedrich Blumenbach (1752–1840), die im Rahmen des Projekts „Johann Friedrich Blumenbach – Online“ bearbeitet wurden. „Johann Friedrich Blumenbach – Online“ ist ein Langzeitprojekt im Rahmen des Akademiensprogramms der Union der deutschen Akademien der Wissenschaften, angesiedelt bei der Akademie der Wissenschaften zu Göttingen.

📄 Datensatz

Alexander von Humboldt in Bern, Unselbständige Schriften
1790–1859

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Inkunabeln JA

📄 Dokumenten 1341 74279

📄 Inkunabeln 2227296

📄 Genre Wissenschaft CC BY-SA 4.0

Das Korpus „A. v. Humboldts Publizistik“ enthält die Digitalausgabe von Alexander von Humboldts „Sämtlichen Schriften“ und entspricht der erweiterten Open-Access-Version der 10-bändigen Druckausgabe (2019) herausgegeben von Oliver Lubrich und Thomas Nehrlisch. Damit wird die gesamte Publizistik Humboldts frei zugänglich gemacht, das heißt die rund 750 Aufsätze, Artikel und Essays, die zu Humboldts Lebzeiten in insgesamt 15 Sprachen an über 420 Orten weltweit in 3.600 Drucken und über 1.200 Publikationsorganen erschienen sind (von Humboldts erster Veröffentlichung 1789 bis zu seinem Tod 1859).

📄 Datensatz

GEI-Digital-2020
1648–1921

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Publikations JA

📄 Publikationen 3935 23797442

📄 Inkunabeln 344219847

📄 Genre Gebrauchsliteratur/Schulbuch CC BY-SA 4.0

Das Korpus GEI Digital besteht aus 5.036 überwiegend deutschsprachigen historischen Schulbüchern die zwischen 1648 und 1921 publiziert wurden.

📄 Datensatz

AEDR: Frühe Neuzeit
1562–1742

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Publikations JA

📄 Inkunabeln 110 17922

📄 Inkunabeln 418882

📄 Genre Gebrauchsliteratur/Lehrerredige CC BY-SA 4.0

Das AEDR-Korpus umfasst vor allem deutschsprachige Funeralschriften aus der Mitte des 16. bis Mitte des 18. Jahrhunderts. Die Auswahl der Schriften entspricht dem Katalog ausgewählter Leichenpredigten der ehemaligen Stadtbibliothek Breslau, der von Rudolf Lenz (bearbeitet von Ralf Berg, Eva-Maria Dickhaut, Sigrid Hubert und Martin Kügel) 1985 herausgegeben wurde. Zwar konzentrieren sich die Provenienzen der Schriften auf den schlesischen Raum, eine bedeutende Anzahl von Schriften ist jedoch auch in weiten Teilen Mitteleuropas veröffentlicht worden.

📄 Datensatz

SBB-Funeralschriften
1643–1659

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Inkunabeln JA

📄 Dokumenten 113 3896

📄 Inkunabeln 89356

📄 Genre Gebrauchsliteratur/Gelegenheitsdruckschrift CC BY-SA 4.0

Das Korpus „SBB-Funeralschriften“ umfasst 112 Werke (1643–1659) von Simon Dach aus dem historischen Druckschriftenbestand der Staatsbibliothek zu Berlin. Die Funeralschriften sind Bestandteil der Sammlung von ca. 20.000 Personalschriften, darunter auch Hochzeitpredigten, Glückwunschsprechliche und andere sogenannte Gelegenheitschriften.

📄 Datensatz

Soldatenbriefe des 18. und 19. Jahrhunderts
1745–1872

🔗 Link Landing Page 🔍 Suche

🔍 Suche

📄 Publikations JA

📄 Dokumenten 178 3119 80023

📄 Genre Gebrauchsliteratur/Brief CC BY-SA 4.0

Das Korpus „Soldatenbriefe“ umfasst 178 Briefe von Offizieren, Unteroffizieren und einfachen Soldaten, adressiert an die Familien in der Heimat. Die Briefe wurden im Zeitraum von 1745 bis 1872 auf Deutsch verfasst; ein deutlicher Schwerpunkt liegt auf Briefen aus den Koalitions- und Befreiungskriegen (1792–1815), dem Deutschen Krieg (1866) und dem Deutsch-Französischen Krieg (1870/71).

📄 Datensatz

Soldatenbriefe im Kontext der DTA-Sammlungen.

3. Das Deutsche Textarchiv (DTA)

Soldatenbriefe des 18. und 19. Jahrhunderts

 Gebrauchsliteratur: Brief	 deu	 1745–1872
 170 Dokumente	 3319 Sätze	 96023 Tokens
 DTABf, TCF	 manually	 ohne Faksimiles
 Soldatenbriefe	 N/A	 CC-BY-SA 4.0

Das Korpus „Soldatenbriefe“ umfasst 170 Briefe von Offizieren, Unteroffizieren und einfachen Soldaten, adressiert an die Familien in der Heimat. Die Briefe wurden im Zeitraum von 1745 bis 1872 auf Deutsch verfasst; ein deutlicher Schwerpunkt liegt auf Briefen aus den Koalitions- und Befreiungskriegen (1792–1815), dem Deutschen Krieg (1866) und dem Deutsch-Französischen Krieg (1870/71).

Stimmen die Schreibentwürfe auch weitgehend überein, so ist die Schreibverfahren der drei Dienstgradgruppen (Offizier, Unteroffizier, Mannschaften bzw. einfacher Soldat) doch sehr verschieden. Das Korpus bietet somit die Möglichkeit, die ganze Breite historischer Alltagsschriftlichkeit an einer einzelnen Briefsorte in den Blick zu nehmen. Das Korpus ist im Rahmen einer sprachhistorischen Untersuchung (Neumann 2019) entstanden und wird dort eingehend beschrieben (vgl. S. 59–67). Die handschriftlichen Brieforiginale stammen aus unterschiedlichen Archiven des gesamten deutschsprachigen Raums. Die Texte wurden buchstabengetreu transkribiert (vgl. S. 72–75).

Suche

Die Korpusbereitstellung ermöglicht komplexe [Suchanfragen \(dstar-Schnittstelle\)](#) auf Grundlage linguistischer Voranalysen (Lemmatisierung, POS-Tagging, orthographische Normierung), Korpusanalysen mit [Wortverlaufscurven](#) und Kollokationsanalysen im zeitlichen Verlauf mit [DiaCollo](#).

Kontakt

- [Dr. Marko Neumann](#) (Transcriber)
- [Deutsches Textarchiv \(DTA\)](#) (annotator, curator, quality-control, IT, support)

Korpus-Website der Soldatenbriefe im DTA.

3. Das Deutsche Textarchiv (DTA)

Zusammenfassung

- Archiv für strukturierte, historische, v. a. deutschsprachige Textsammlungen
- Bereitstellung von Texten aus dem Zeitraum von ca. 1600 bis 1920
- sorgfältige Auswahl der Texte und Ausgaben (vollständige Werke)
- hohe Erfassungsgenauigkeit
- strukturelle und linguistische Erschließung der Textdaten
- Verlässlichkeit der Metadaten und Verwendung von Standards (DTABf, CMDI, etc.)
- Interoperabilität

4. Infrastruktur am Zentrum Sprache

4.1 Das Zentrum Sprache

- zentrales Vorhaben: Digitales Wörterbuch der deutschen Sprache (**DWDS**)
- Wortinformationssystem, das lexikalisches Wissen mit authentischen Textbelegen verbindet
 - => es werden umfangreiche Textkorpora verwendet
- Aufbereitung der Texte durch **(computer-)linguistische Methoden**
- Bereitstellung einer (korpusübergreifenden) **Suche** (inkl. Anbindung an Thesauri)
- Bereitstellung von Tools zur **Korpusanalyse** (Wortverlaufskurven, DiaCollo)

4. Infrastruktur am Zentrum Sprache

4.2 Korpussuche I

Suchanfrage	Treffer
Krankenhaus	0
Krankenhaus germanet	Lazareth, Spital, Sbital, Hospital, Spiethal

Query: <https://www.dwds.de/r/?q=&corpus=soldatenbriefe>.

4. Infrastruktur am Zentrum Sprache

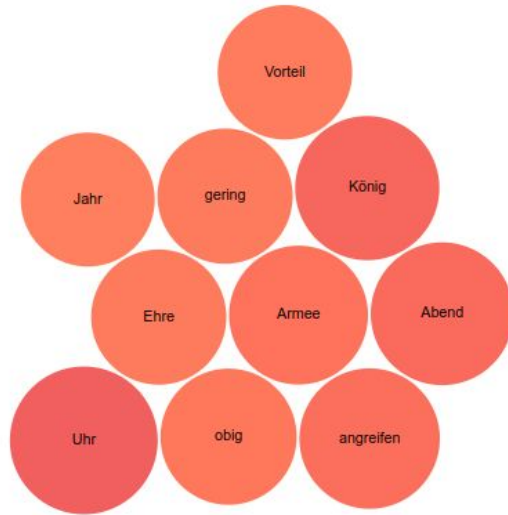
4.3 Korpussuche II

Suchanfrage	Treffer (Anzahl)
Alkohol germanet	102
count(Alkohol germanet) #by[\$w, role] #desc_count	13 Wasser Mannschaften 9 Wein Mannschaften 8 Bier Mannschaften 6 Wasser Offizier 5 Schnaps Mannschaften 5 ale Unteroffizier 4 Wein Unteroffizier

Query: <https://www.dwds.de/r/?q=&corpus=dtaxl>.

4. Infrastruktur am Zentrum Sprache

4.4 Korpusanalyse mit DiaCollo



1750



1810



1860

Query: <https://kaskade.dwds.de/dstar/soldatenbriefe/diacollo/?query=König>.

5. Nachnutzung durch correspSearch

The screenshot shows the correspSearch interface. At the top, there is a navigation bar with the logo 'correspSearch' and menu items: Start, Suche, Daten, Über correspSearch, Mitmachen, API, quoteSalute, and Deutsch. Below this is a search bar with the text 'Briefpartner:in' and a dropdown arrow. To the right of the search bar is the text 'Erweiterte Suche'. Below the search bar, there is a filter for 'Soldatenbriefe des 18. und 19. Jahrhunderts'. The main content area shows '1-10 von 170 Ergebnissen' and a pagination control with page numbers 1, 2, 3, 4, and an ellipsis. Two search results are visible:

- 19. Juni 1745**
Königgretz
Nikolaus Binn → Ehefrau
In: Neumann, Marko. Soldatenbriefe des 18. und 19. Jahrhunderts: Untersuchungen zu Syntax und Textstruktur in der Alltagsschiflichkeit unterschiedlicher militärischer Dienstgrade. Germanistische Bibliothek, Band 68. Heidelberg: Universitätsverlag Winter, 2019. Digitaler Korpus unter <https://www.dwds.de/d/korpora/soldatenbriefe>
Zum Brief
- 21. September 1756**
Außig
Zetlingen
Christian Arnholtz → Bruder
In: Neumann, Marko. Soldatenbriefe des 18. und 19. Jahrhunderts: Untersuchungen zu Syntax und Textstruktur in der Alltagsschiflichkeit unterschiedlicher militärischer Dienstgrade. Germanistische Bibliothek, Band 68. Heidelberg: Universitätsverlag Winter, 2019. Digitaler Korpus unter <https://www.dwds.de/d/korpora/soldatenbriefe>

On the right side, there is a sidebar with a 'Zeitraum' section showing a bar chart of results from 1745 to 1872. Below the chart is a list of years with counts and expand/collapse icons:

Year	Count
1745	1
1756	2
1757	3
1758	3
1759	1

Below this is a 'Personen' section with a list of names and counts:

Name	Count
Pogge, Carl	3
Callot, Eduard Ferdinand von	3

<https://correspsearch.net/de/suche.html?c=>

<https://raw.githubusercontent.com/deustchestextarchiv/soldatenbriefe/main/scripts/CMIF.xml>

Zusammenfassung

DTA:

- Beratung zur Aufnahme von historischen, deutschsprachigen Texten
- standardkonforme Annotation (TEI, DTABf) bestenfalls unter Verwendung von Normdaten
- Kuration, Integration und **Bereitstellung der Sammlung** gemäß der FAIR-Prinzipien unter einer möglichst offenen Lizenz (CC BY-SA 4.0)

Zentrum Sprache:

- linguistische Aufbereitung: (übergreifende) **Suche, Korpusanalyse**
- Bereitstellung von **Schnittstellen** zur Nachnutzung der Daten

Text+:

- konsortienweite Sichtbarkeit der Sammlung inkl. Provenienz
- übergreifende Suchmöglichkeit

Fragen, Anregungen und Kontakt

Zu den Soldatenbriefen:

marko.neumann@germanistik.uni-giessen.de

Zum Deutschen Textarchiv, zu historischen Korpora etc.:

marius.hug@bbaw.de