



Überblick: Semantische Korpusanalyse mit dem WdK-Explorer

Ben Heuwing

Institut für Informationswissenschaft & Sprachtechnologie

Universität Hildesheim

Projektfokus

Entwicklung digitaler Werkzeuge zur Analyse großer Quellenbestände, die mit hermeneutischen Methoden kaum zu durchdringen sind.

Instrumente sollen der Semantik der Bildungsmedien gerecht werden und 'qualitative' Strukturen aufdecken:

- intertextuelle Verknüpfungen
- thematische Cluster
- semantische Felder

Anwendung auf geschichtswissenschaftliche Fragestellungen



Korpus

Historische Schulbücher

- 3803 digitalisierte Schulbücher
- Geschichte, Geografie, Realienkunde
- <http://gei-digital.gei.de>

Insgesamt 799260 Text-Seiten

Umfangreiche Metadaten



Letzte Importe

23. Oktober 2015
Geographie für Kinder

Grundriß der Europäischen Staatengeschichte, in Verbindung der Erdbeschreibung und Staatskunde

Hieronymi Freyers Paed. Reg. Hal. Insp. Nähere Einleitung zur Universal-Historie

07. Oktober 2015
M. Georg Christian Raff's, ordentlichen Lehrers der Geschichte und Geographie auf dem Lyceum zu Göttingen, Geographie für Kinder

27. August 2015
Schul-Lesebuch

Sie befinden sich hier: [Startseite](#)

Verfügbare Sammlungen



Fibeln vor 1871 (54)

Fibeln oder auch ABC-Bücher waren meist bebilderte, häufig aufwändig illustrierte Erstlesebücher, die als Schulbücher zum Zweck des Erlernens des Lesens eingesetzt wurden. Die Sammlung umfasst Erstlesebücher aus Deutschland vom 15. / 16. Jahrhundert an bis in das Jahr 1870.



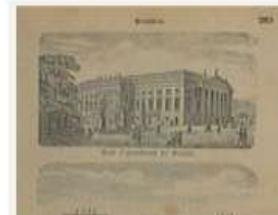
Geographieatlanten (119)

Geographieatlanten stellen als didaktisch aufbereitete, systematisch angeordnete und gebundene Sammlungen von Karten eine Sonderform der Atlanten dar und wurden hauptsächlich im Geographieunterricht eingesetzt. In der vorliegenden digitalen Sammlung finden Sie Geographieatlanten, die in der Zeit zwischen dem 17. Jahrhundert und dem Ende des Kaiserreiches im Jahr 1918 erschienen sind.



Geographieschulbücher Kaiserreich (956)

Für den Geographieunterricht wurden Schulbücher verwendet, wie sie in der vorliegenden Sammlung zu finden sind. Der zeitliche Rahmen umfasst die Zeit des Kaiserreiches zwischen 1871 und 1918.



Geographieschulbücher vor 1871 (265)

Die Schulbücher dieser digitalen Sammlung wurden in der Zeit vom 17. Jahrhundert bis in das Jahr 1870 für den Geographieunterricht verwendet.



Geschichtsatlanten (53)

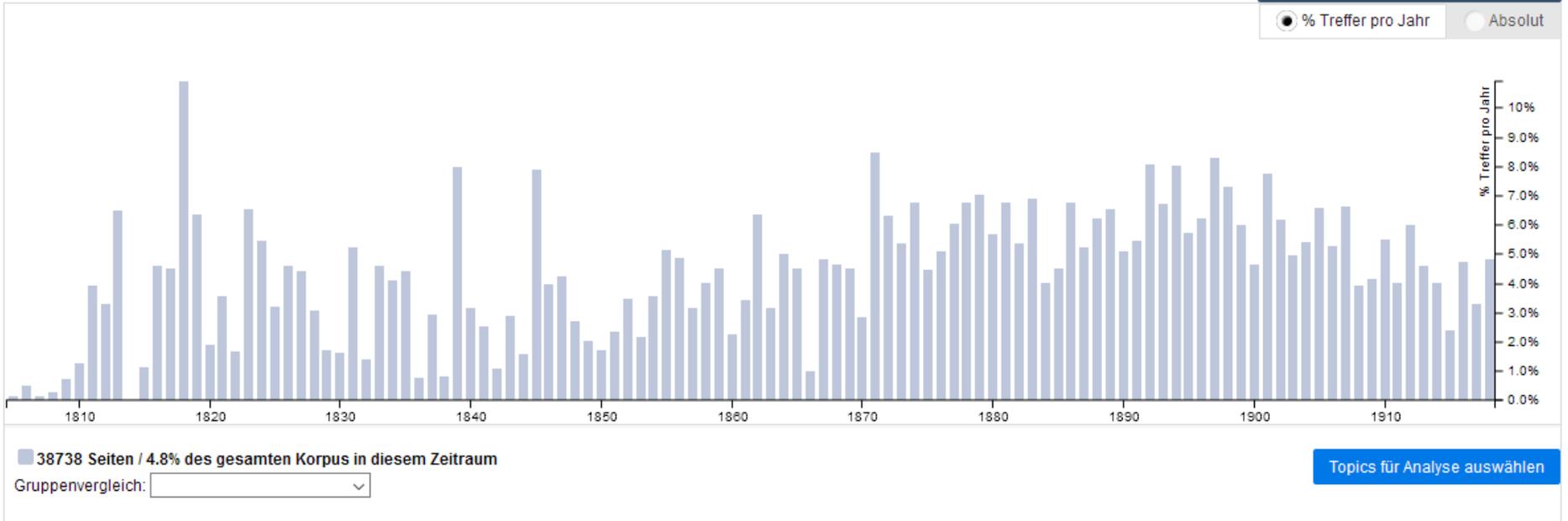
napoleon

Suchen

Aktuelle Auswahl: [Anfrage:napoleon](#) [min. 50 Worte, min. 3 Sätze, geringer Anteil Sonderzeichen/Wort](#)

Diagramm für Aktuelle Auswahl

% Treffer pro Jahr Absolut



Vergleichen: Kategorie x Kategorie

Ergebnisse filtern:

Anzeige: [Seitenzahl](#) | % der Ergebnisse

Junkfilter

Auflage

Erscheinungsjahr

1800 : 1920 OK

1800 - 1809	0%
1810 - 1819	2%
1820 - 1829	2%
1830 - 1839	3%
1840 - 1849	2%
1850 - 1859	3%
1860 - 1869	5%
1870 - 1879	10%
1880 - 1889	12%
1890 - 1899	16%
1900 - 1909	19%
1910 - 1919	21%

Von 1800 bis 1920 100%

Sammlung

Inhalt Raum/Thema

Inhalt: Zeit

Schultyp_Allg_wdk

Schultyp_wdk

Seiten

Gruppirt nach Werk

38738 Seiten

Treffer pro Seite: Treffer bis 10 von 38738 [weiter»](#) [»»](#)

Sortieren: [Relevanz zu Anfragetermen](#) [Erscheinungsjahr](#) [Seitenzahl](#)

1. Der geschichtliche Unterricht in der Volksschule ▾ - S. 97

1910 - München : Kellerer

Autor: Haberl, Johann

Sammlung: Kaiserreich Geschichtsschulbuecher

Schulbuchtyp_wdk: Lehrerbuch

Schultyp_wdk: Volksschule, Landschule

Schultyp_Allg_wdk: Niedere Lehranstalten

Suchanfragen

kaiser wilhelm Alle Terme Suchen

➔ **32744** Seiten

kaiser wilhelm min. 1 Term Suchen

➔ **191578** Seiten

Kombination der Anfrageterme:

Trefferdokumente müssen **Alle Terme** enthalten (Standardeinstellung - Verknüpfung mit UND)

Trefferdokumente müssen **mindestens einen der Terme** enthalten (Verknüpfung mit *ODER*)

Einfaches Stemming wird auf die Suchanfrage angewendet

Regierungen → *regierung* (findet *regierungen* und *regierung*)

regierte → *regiert* (findet Seiten mit *regierte* und *regiert*, aber nicht *regieren*)

Kaiser → *kais* (findet *Kaiser*, aber nicht *Kaiserin*)

Anfrage-Syntax: Erweiterte Möglichkeiten

kaiser	Alle Seiten mit <i>kaiser</i> oder einer Flektion (über einfaches Stemming zu kais) im Text. Groß-Kleinschreibung spielt keine Rolle
kai*	Alle Seiten mit kaiser oder kai (z.B. im Hamburger Hafen) und mit längeren Termen
kaiser AND könig	Alle Seiten sowohl mit dem Term <i>kaiser</i> als auch <i>könig</i> , entspricht der Anfrage <i>kaiser könig</i>
Kaiser OR könig	Alle Seiten entweder mit dem Term kaiser oder könig oder mit beiden
Kaiser -könig	Alle Seiten mit dem Term kaiser aber ohne könig
(Kaiser OR könig) AND Wilhelm	Kombinationen aus AND / OR können mit Klammern kombiniert werden
"deutschen volkes"	Alle Seiten, auf denen deutsch und volk direkt hintereinander auftraten (Phrasensuche) – Achtung: Es werden auch Wort-Flektionen berücksichtigt (stemming): Es wird also nicht nach der exakten Formulierung gesucht!
"kaisertum Glanzzeit"~2	Alle Seiten, auf denen eine Phrase wie „Kaisertum seine Glanzzeiten“ oder “Kaisertums in seiner Glanzzeit“ auftritt: Die Begriffe dürfen 2x verändert, d.h. umgestellt oder zusätzliche Begriffe eingefügt werden
goobi_CatalogIDSource:PPN564100781	Alle Seiten aus dem Werk mit der PPN PPN564100781 (Für Suche nach exaktem Wert)
goobi_TitleDocMain:Hilfsbuch	Alle Seiten aus Werken, die Hilfsbuch im Titel haben (für indexierte Felder text, text_normalized)

<https://cwiki.apache.org/confluence/display/solr/The+Extended+DisMax+Query+Parser>

Alle aktiven Anfrageterme und Filter

werden unterhalb der Anfrage angezeigt

regierung min. 1 Term Suchen ?

Aktuelle Auswahl: [Anfrage \(min. 1 Term\): regierung](#) | [Jahr der Erstauflage \(WdK\) : 1900 bis 1909](#) | [min. 50 Worte, min. 3 Sätze, geringer Anteil Sonderzeichen/Wort](#) ✕ Auswahl zurücksetzen



Alle aktiven Teile der Anfrage löschen / Alle Dokumente zurückgeben

Vorhandene Metadaten und darauf basierende Filter

Anzahl Seiten

Bibliographische Angaben zu jedem Werk: Titel, Erscheinungsort, Verlag, Herausgeber

Metadaten aus OCR-Prozess, aus Bibliothek des GEI und aus dem Projekt WdK

Beispiele:

- Verlag / Verlag (gesäubert WdK): Ursprünglicher Wert und einmal Schreibweisen normalisiert
- Erscheinungsort / Ort_erste_Angabe: Reduziert auf erste Ortsnennung
- Schulform_opac: Extrahiert aus Opac-Verschlagwortung
- Erscheinungsjahr: Jahrzehnt oder beliebige Zeitspanne
- Inhalt Raum/Thema: Forschungsbezug WdK

→ **Cheat-Sheet mit Filternamen**



▼ Erscheinungsjahr	
1800 - 1809	1401
1810 - 1819	1632
1820 - 1829	2988
1830 - 1839	4123
1840 - 1849	3603
1850 - 1859	5292
1860 - 1869	7074
1870 - 1879	11926
1880 - 1889	15245
1890 - 1899	17965
1900 - 1909	22068
1910 - 1919	25020
Vor 1800	1270
Von 1800 bis 1920	118337

▼ Sammlung	
<input type="checkbox"/> Kaiserreich Geschichtsschulbuecher	72106
<input type="checkbox"/> Geschichtsschulbuecher vor 1871	19996
<input type="checkbox"/> Geographieschulbuecher Kaiserreich	9281

Spezielle Filter

- ▶ Volltext
- ▶ Extrahierte Personennamen
- ▼ Extrahierte Ortsnamen
 - england 24%
 - deutschland 23%
 - frankreich 20%
 - afrika 17%
 - spani 17%
 - europa 16%

Filter für häufige Terme und Namen aktivieren

Text- und Namensfacetten aktiv

Anzeigen:

Häufige Namen und Terme:

Rechenintensiv, müssen zunächst aktiviert werden.

▼ Junkfilter

- min. 50 Worte, min. 3 Sätze, geringer Anteil Sonderzeichen/Wort
- Mindestens 5 Worteinheiten/Satz
- Medientypen filtern

Junkfilter:

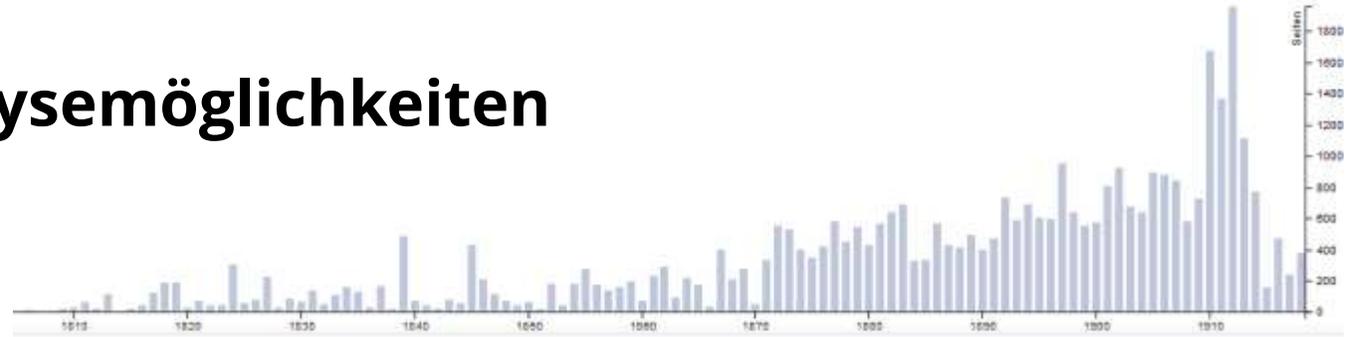
Standardmäßig aktiviert, Seiten mit sehr wenig korrekt erkanntem Text herausfiltern

1. **Kleine vaterländische Geschichte** ▼ - S. 83
1883 - Langensalza : Beyer
Autor: Wolf, Carl
Hrsg.: ,
Auflage: 2
Sammlung: Kaiserreich Geschichtsschulbuecher
Schulbuchtyp_wdk: Schülerbuch

Filter auf Ergebnisse aus einzelnen Werk

Analyse im Zeitverlauf - Analysemöglichkeiten

Seitenzahl im Zeitverlauf



Im Gruppenvergleich oder Vergleich mit vorheriger Anfrage

Gruppenvergleich

Vergleich mit vorheriger Anfrage

Topic Intensität für ausgewählte Topics anzeigen

Oder ein einzelnes Topic im Vergleich

Topics für Analyse auswählen

Optionen

Gleitender Durchschnitt über 5 Jahre, um Trends zu erkennen

Wenn Anfrage oder Filter aktiv: Normalisiert an Seitenzahl im jeweiligen Jahr um Vergleiche zwischen Jahren zu ermöglichen

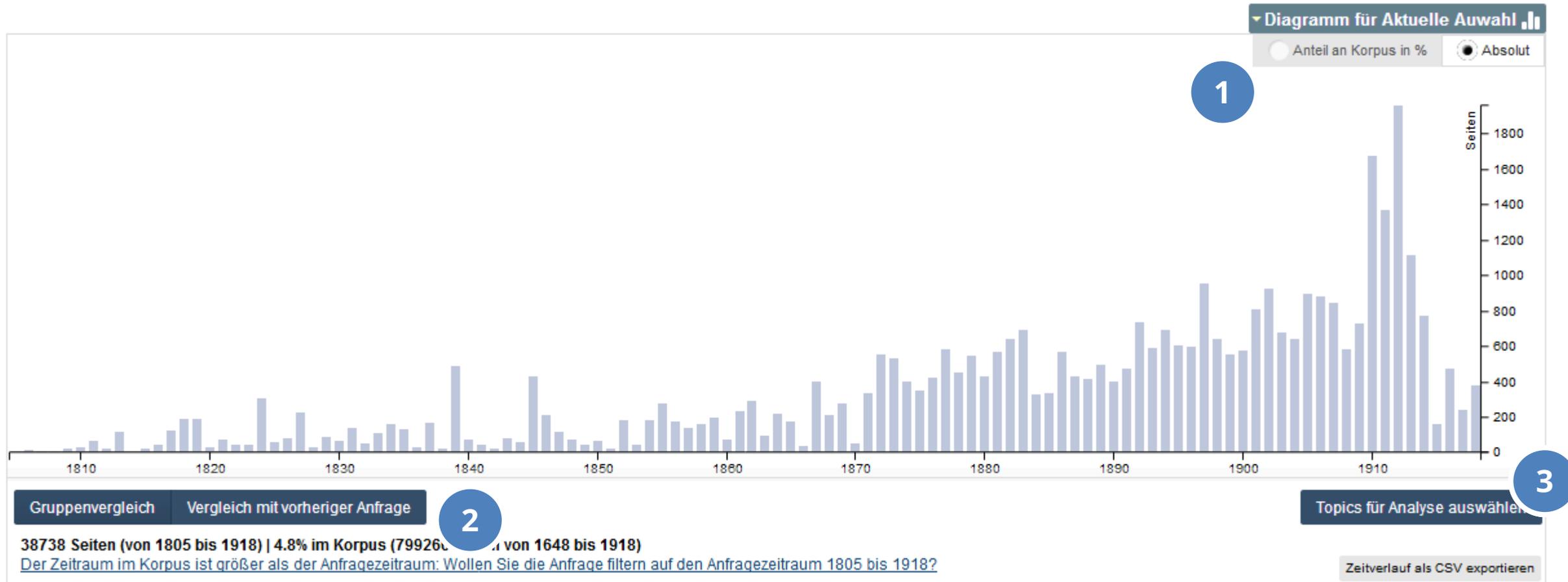
1 Jahr

5 Jahre

Anteil an Korpus in %

Absolut

Treffer (Seitenzahl) in aktueller Auswahl im Zeitverlauf

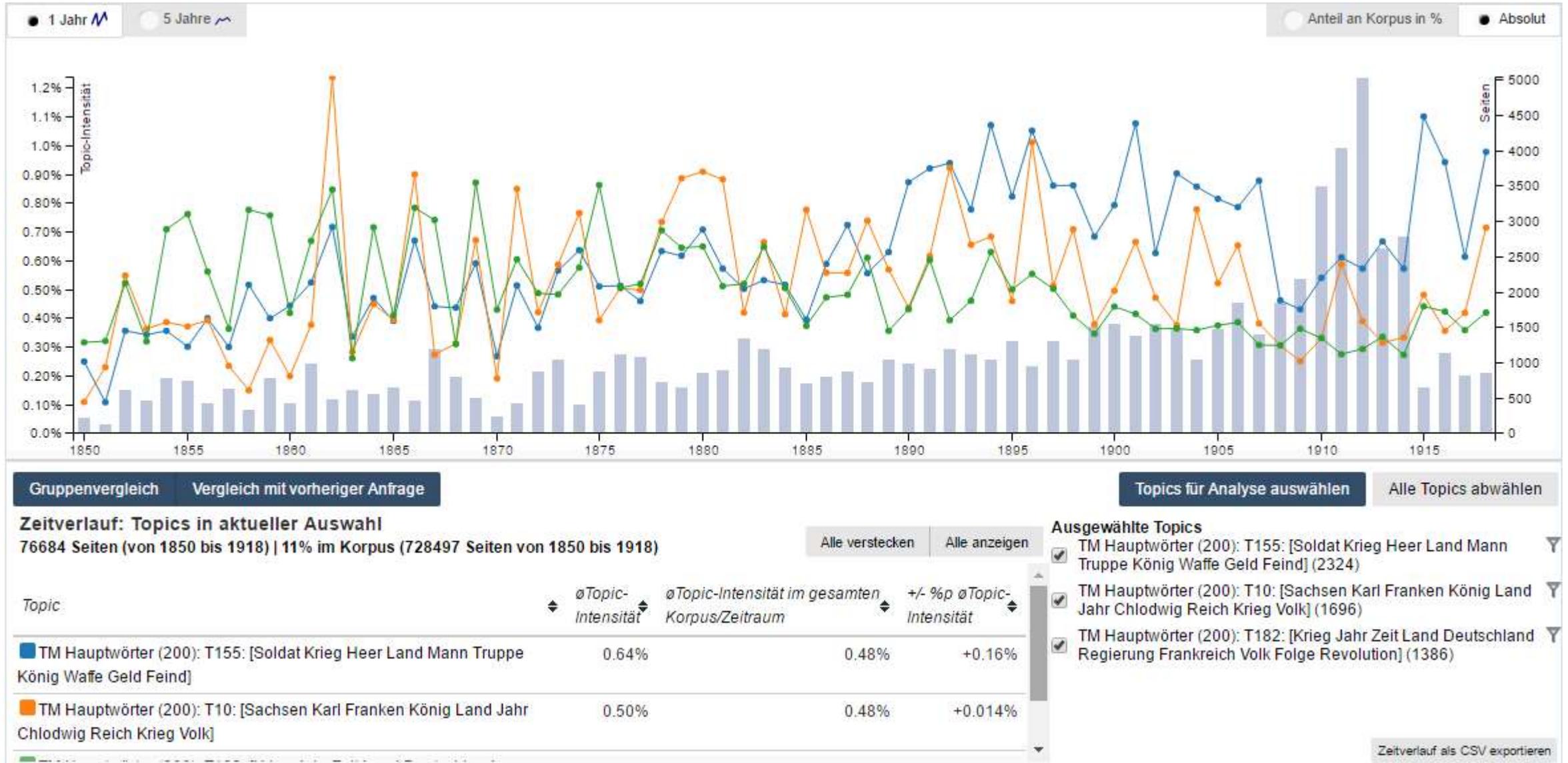


1) Absolute Anzahl / Anteil in dem jeweiligen Jahr

2) Gruppenvergleich oder Vergleichsanfrage auswählen

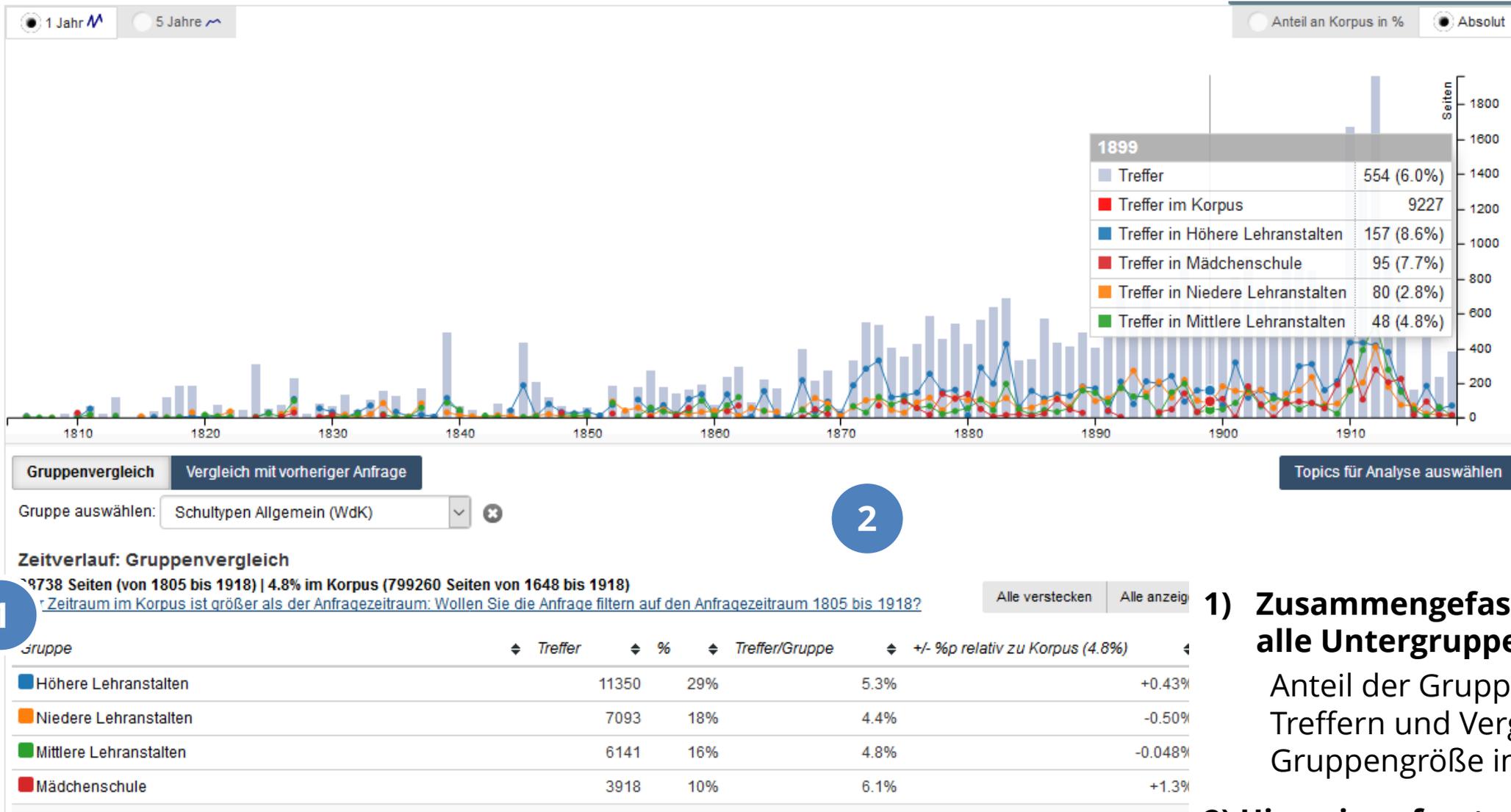
3) Topics für Analyse auswählen

Vergleich von Topics im Zeitverlauf



Beispiel: Kinder (Suchanfrage) und Krieg (Topics)

Gruppenvergleich im Zeitverlauf



1

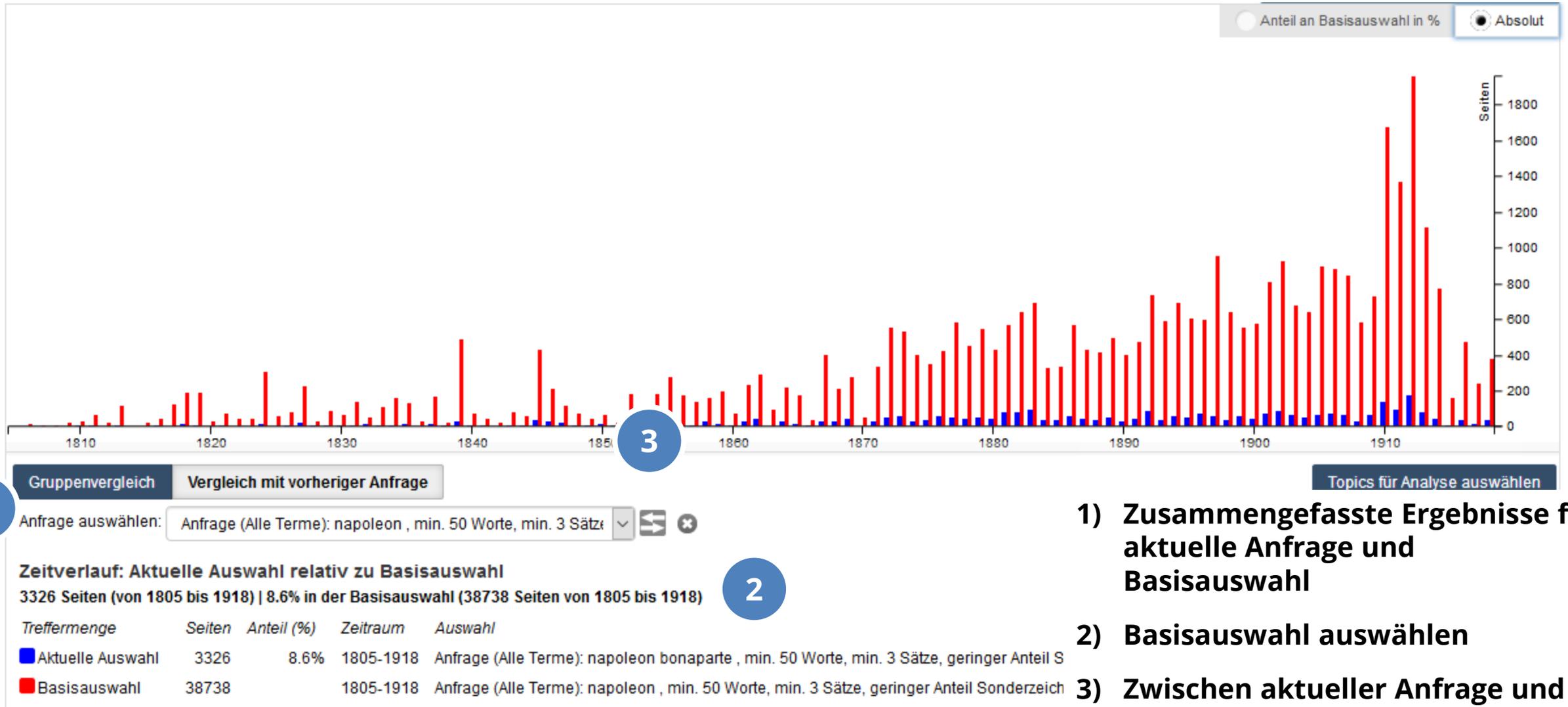
2

1) Zusammengefasste Ergebnisse für alle Untergruppen:

Anteil der Gruppe an aktuellen Treffern und Vergleich mit Gruppengröße in Gesamtkorpus

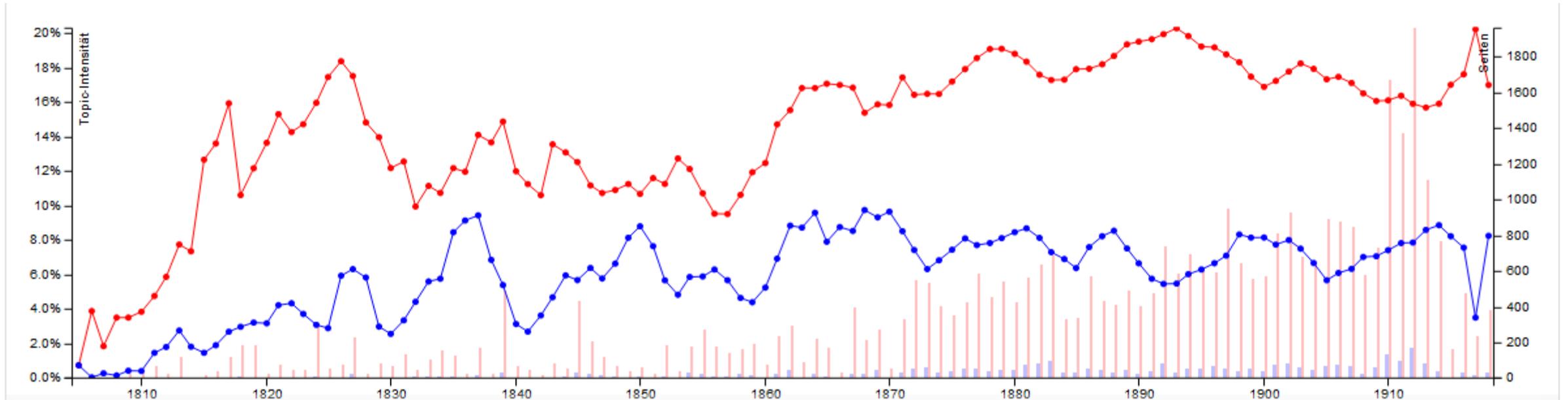
2) Hinweis auf unterschiedliche Zeitabschnitte zwischen Anfrage und Korpus durch Auswahl

Zeitverlauf aktuelle Auswahl im Vergleich zu vorheriger Anfrage als Basisauswahl



- 1) Zusammengefasste Ergebnisse für aktuelle Anfrage und Basisauswahl
- 2) Basisauswahl auswählen
- 3) Zwischen aktueller Anfrage und Vergleichsbasis wechseln

Topic im Zeitverlauf für aktuelle Auswahl im Vergleich zu vorheriger Anfrage als Basisauswahl



Gruppenvergleich | **Vergleich mit vorheriger Anfrage** | Topics für Analyse auswählen | Alle Topics abwählen

Anfrage auswählen: Anfrage (Alle Terme): napoleon , min. 50 Worte, min. 3 Sätze | Anzeigen: T28: [Schlacht Heer Feind Mann Armee N ...]

Zeitverlauf: Aktuelle Auswahl relativ zu Basisauswahl mit Topic
 3326 Seiten (von 1805 bis 1918) | 8.6% in der Basisauswahl (38738 Seiten von 1805 bis 1918)

Topic	∅ Topic-Intensität	Seiten	Anteil (%)	Zeitraum	Auswahl
■ TM Hauptwörter (50): T28: [Schlacht Heer Feind Mann Armee Napoleon Franzose General Truppe Preußen] in der Basisauswahl	17%	38738		1805-1918	Anfrage (Alle Terme): napoleon , min. 50 Worte, min. 3 Sätze, geringer Anteil Sonderzeichen/Wort
■ TM Hauptwörter (50): T28: [Schlacht Heer Feind Mann Armee Napoleon Franzose General Truppe Preußen]	7.3%	3326	8.6%	1805-1918	Anfrage (Alle Terme): napoleon , min. 50 Worte, min. 3 Sätze

Ausgewählte Topics

- TM Hauptwörter (50): T28: [Schlacht Heer Feind Mann Armee Napoleon Franzose General Truppe Preußen] (701)

Relative Ansicht zeigt Differenz zwischen Topic in den beiden Mengen

Ähnliche Ergebnisse zu einzelnen Treffern: More Like This

Zu jedem Dokument in WdK-Explore können Sie jeweils ähnliche Dokumente anzeigen lassen

Ähnlichkeit auf der Basis der in dem Dokument enthaltenen Terme oder eingeschränkt auf die enthaltenen Personen und Ortsnamen

TM Hauptwörter (50): [T45: [Zeit Mensch Leben Kunst Sprache Wissenschaft Natur Wort Geist Lehrer], T10: [Volk König Mann L deutsch Geschichte]]
TM Hauptwörter (100): [T92: [Mensch Leben Natur Arbeit Zeit Ding Geist Welt Art Seele], T43: [Zeit Volk Jahrhundert Geschichte Himme], T52: [Mensch Leben Volk Gott Geist Zeit Religion Mann Glaube Herz]]
TM Hauptwörter (200): [T127: [Volk Sprache Land Zeit Sitte Kultur Bildung Geschichte Bewohner Stamm], T136: [Leben Mensch Zeit], T175: [Mensch Leben Natur Körper Seele Tier Tiere Arbeit Erde Pflanze]]
[Seite auf gei-digital.de](#) Ähnliche Ergebnisse: [nach allen Termen](#) | [nach enthaltenen Personen- & Ortsnamen](#)

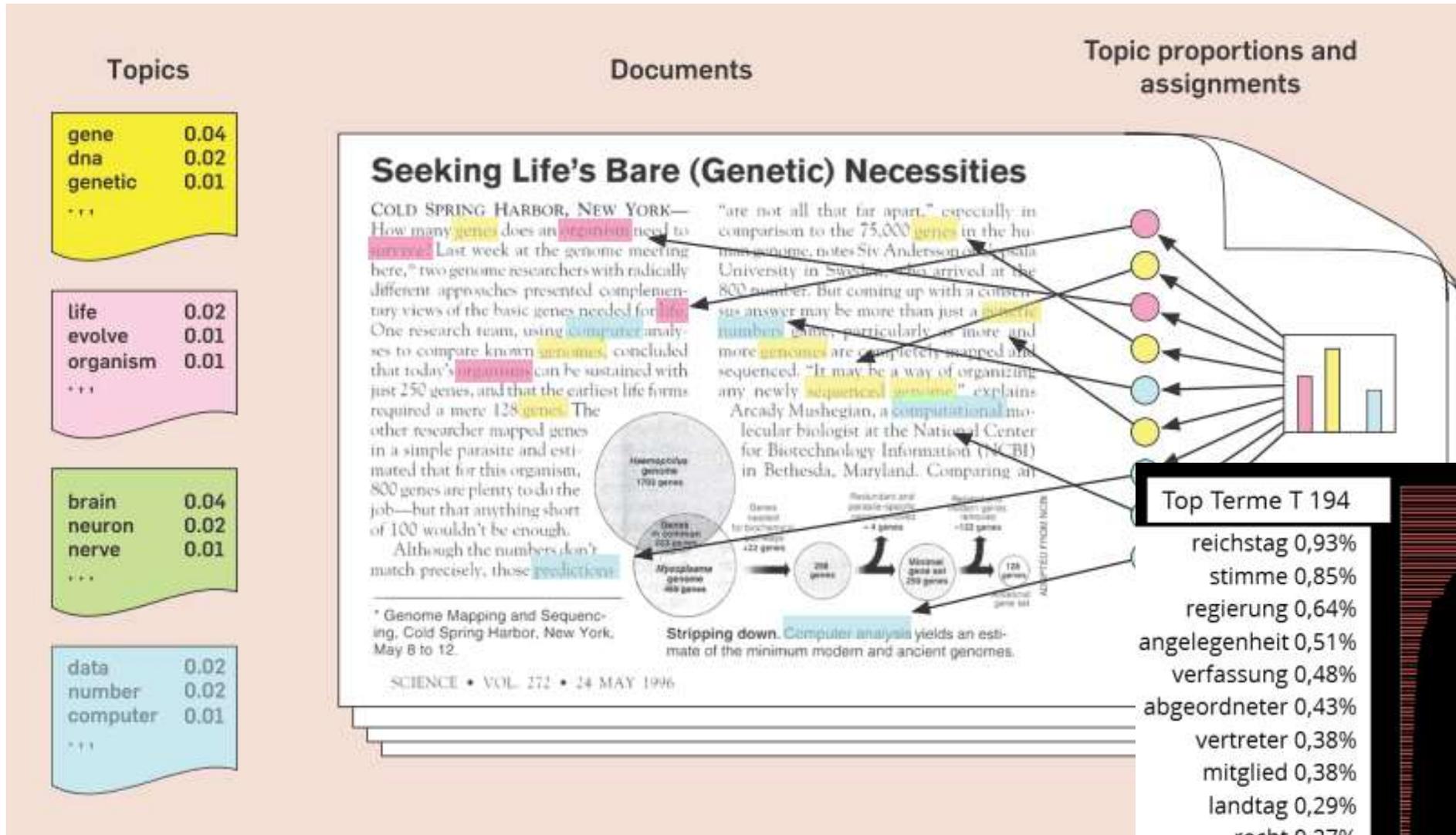


Ähnliche Ergebnisse

Maximale Anzahl Ergebnisse:

Ähnliche Dokumente basierend auf den Feldern **Volltext**

Topic Models



Automatisiertes Verfahren, kein manuell erstelltes Trainingsmaterial notwendig

Wichtigste Terme als Name eines Topics

Mehrere Topics pro Dokument

Blei, David M., 'Probabilistic Topic Models', *Communications of the ACM*, 55 (2012), 77-84

Top Terme T 194

- reichstag 0,93%
- stimme 0,85%
- regierung 0,64%
- angelegenheit 0,51%
- verfassung 0,48%
- abgeordneter 0,43%
- vertreter 0,38%
- mitglied 0,38%
- landtag 0,29%
- recht 0,27%
- bundesstaat 0,22%
- stand 0,22%
- ...

Topic Modeling: Geeignete Einführungen & Kritik

Blei, D. M. (2012). Topic Modeling and Digital Humanities. *Journal of Digital Humanities*, 2(1). Abgerufen August 21, 2014, von <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/>

Weingart, S. (2012, Juli 25). Topic Modeling for Humanists: A Guided Tour. *the scottbot irregular*. Abgerufen September 13, 2014, von <http://www.scottbot.net/HIAL/?p=19113>

tedunderwood. (2012). Topic modeling made just simple enough. *The Stone and the Shell*. Abgerufen August 21, 2014, von <http://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/>

Schmidt, B. M. (2013). Words Alone: Dismantling Topic Models in the Humanities. *Journal of Digital Humanities*. Abgerufen April 7, 2016, von <http://journalofdigitalhumanities.org/2-1/words-alone-by-benjamin-m-schmidt/>

Topic-Modeling im Projekt

Vorverarbeitung

- OCR-Nachkorrektur
- Lemmatisierung -> Reduktion auf Grundformen
- Nur großgeschriebene Lemmata -> Substantive

3 Topicmodelle verfügbar

- 50 Topics
- 100 Topics
- 200 Topics

TM 50 Topics

1. [Gott Mensch Herr Herz Leben Wort Welt Himmel Tag Hand]
2. [Reich Zeit Staat Volk Deutschland Jahrhundert Land Macht deutsch Geschichte]
13. [Athen Stadt Athener Sparta Spartaner Griechenland Krieg Perser Flotte König]
36. [Handel Industrie Land Ackerbau Fabrik Stadt Deutschland Mill Viehzucht Gewerbe]

TM 200 Topics

- 1: [Kind Lehrer Schüler Unterricht Schule Frage Stoff Aufgabe Zeit Geschichte]
2. [Staat Zeit Volk Deutschland Leben Reich Jahrhundert Macht Entwicklung Gebiet]
20. [Preußen Bund Staat König Regierung Deutschland Verfassung Frankfurt Reichstag Bundestag]
62. [Papst Kaiser Iii Vii Gregor Heinrich Rom Friedrich Italien Jahr]

Filtern mit Topics

Topic-Modelle auch als Filter verfügbar

Jeder Seite werden dafür die wichtigsten Topics zugeordnet

Filtern auf ein Topic und gleichzeitige Analyse dieses Topics nicht sinnvoll!

▶	TM Hauptwörter (50)	
▶	TM Hauptwörter (100)	
▼	TM Hauptwörter (200)	
	T9: [Frieden Napoleon Krieg Kaiser	5984
<input type="checkbox"/>	Frankreich Friede Preußen Rußland	
	Jahr Franz]	
	T21: [Napoleon Bluch Heer General	4809
<input type="checkbox"/>	Preußen Franzose Schlacht Armee	
	Mann Wellington]	

Topics auswählen für Analyse

Topics für Analyse auswählen

Topics auswählen ✕

Auswahl:
 TM Hauptwörter (50): T35: [Preußen Königreich Bayern Sachsen Staat Hannover Baden König Provinz Land]

Auswahl übernehmen Abbrechen

Filter: Alle abwählen

#	Name	Treffer	
34	T34: [Krieg Frankreich England Deutschland Preußen Frieden Rußland Napoleon Kaiser Jahr]	8169	<input type="checkbox"/>
10	T10: [Volk König Mann Leben Zeit Land Mensch Krieg Feind Vaterland]	3199	<input type="checkbox"/>
2	T2: [Schweden Friedrich Heer Schlacht Sachsen König Gustav Kaiser Krieg Schlesien]	820	<input type="checkbox"/>
14	T14: [Athen Stadt Athener Sparta Spartaner Griechenland Krieg Perser Flotte König]	74	<input type="checkbox"/>
23	T23: [Rom Römer Krieg Italien Stadt Jahr Heer König Rmer Hannibal]	54	<input type="checkbox"/>
20	T20: [Rom Jahr Cäsar Senat Kaiser Pompejus Antonius Tod Krieg Sohn]	37	<input type="checkbox"/>

▾ TM Hauptwörter (50) - 6 1

▸ TM Hauptwörter (100) - 11

▸ TM Hauptwörter (200) - 19

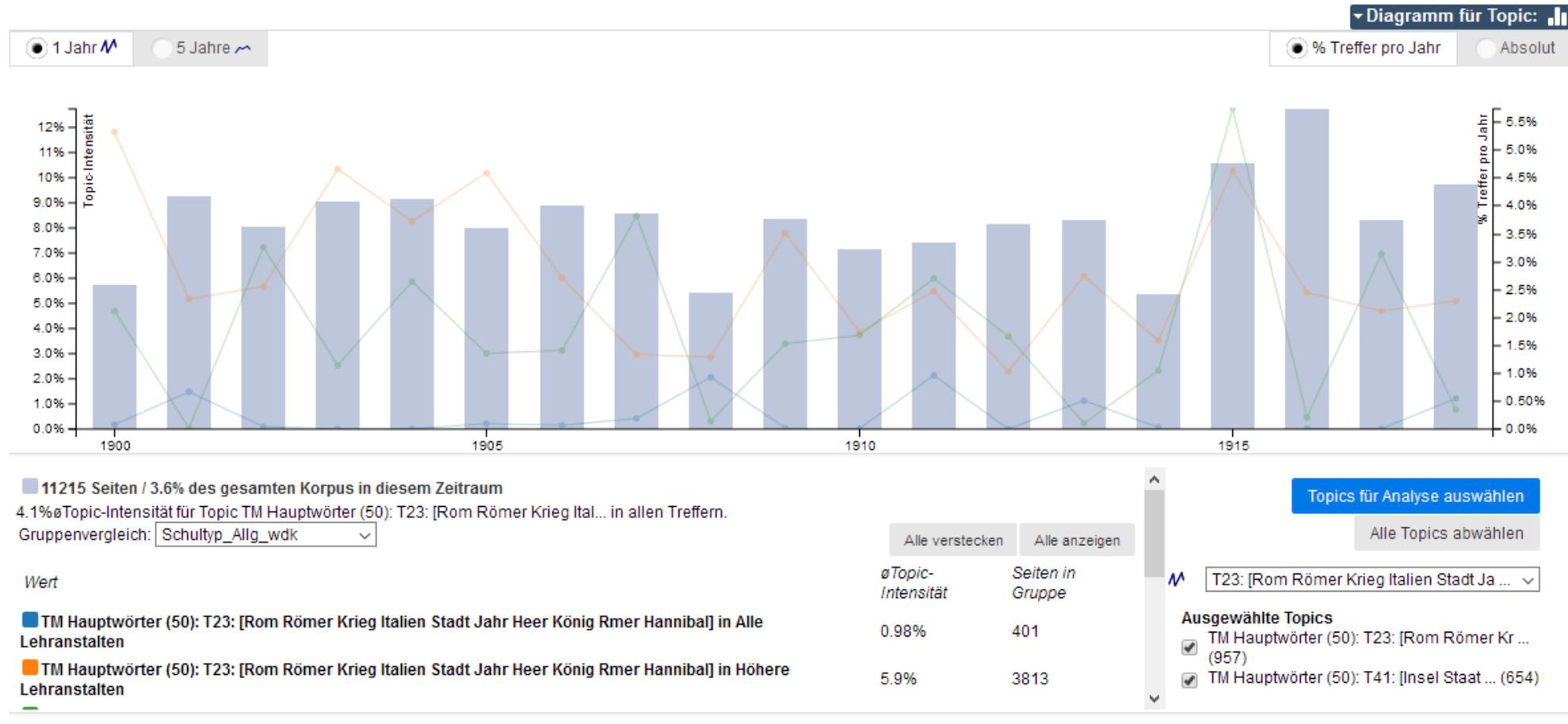
Sortiert nach Häufigkeit in aktuellen Treffern

Filtermöglichkeit nach Topic-Term

Weitere Möglichkeit: Suchterm mit Topic ersetzen

1. Suchterme eingeben
2. Topics werden nach Häufigkeit zu Suchterm sortiert angezeigt
3. Geeignete Topics für Analyse auswählen oder danach filtern
4. Suchterm wieder abwählen (!)

Analyse mit Topics im Diagramm



Mittlere Topic-Intensität im jeweiligen Zeitraum

Mehrere Topics *oder* Untergruppen in einer Kategorie ein Topic

Alle Topics über Gruppen vergleichen

z.B. wichtigste Topics zu „napoleon“ nach Jahrzehnt

Vollständiges Topic Model (am Besten in Zeilen)

Z.B. häufigstes Topic in verschiedenen Zeiträumen,

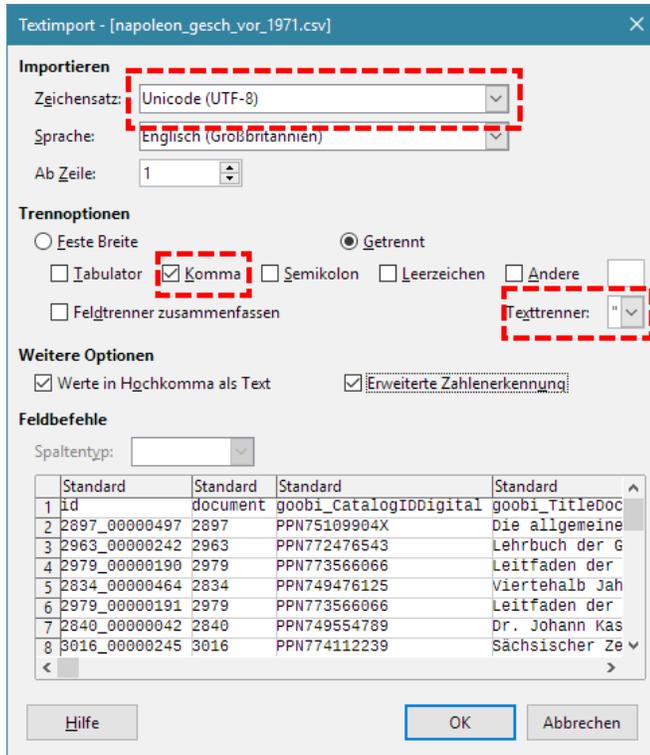
Ausreißer-Topics

Vergleiche: gruppiert nach: Relative Werte Top-Abweichungen anzeigen:

Häufigkeiten für *TM Hauptwörter (200)* im Vergleich über *Erscheinungsjahr n. Jahrzehnten*

Erscheinungsjahr		Gesamt		1860 - 1869		1870 - 1879		1880 - 1889	
#	TM Hauptwörter (200)	Anzahl	Anteil	Anzahl	+/-	Anzahl	+/-	Anzahl	+/-
<i>Top-3-Aufsteiger (nur **):</i>						1. T9: [Frieden Napoleon Krieg...: +9% 2. T182: [Krieg Jahr Zeit Land...: +5% 3. T73: [König Paris Parlament...: +4%		1. T9: [Frieden Napoleon Krieg...: +5% 2. T73: [König Paris Parlament...: +3% 3. T197: [Italien Mailand Stadt Rom...: +3%	
<i>Top-3-Absteiger (nur **):</i>						1. T54: [Staat Zeit Volk Deutschland...: -5% 2. T71: [Deutschland Krieg Preußen...: -5% 3. T35: [König Bismarck Wilhelm...: -5%		1. T54: [Staat Zeit Volk Deutschland...: -3% 2. T183: [Kind Lehrer Schüler...: -3% 3. T65: [König Herr Soldat Offizier...: -1%	
<i>Seitenzahl pro Erscheinungsjahr:</i>		25249		3	0%	3279	12%	4058	16%
1	T9: [Frieden Napoleon Krieg Kaiser Frankreich Friede Preußen Rußland Jahr Franz]	7685	30%	3	+70%	1292	+9%**	1434	+5%**
2	T21: [Napoleon Bluch Heer General Preußen Franzose Schlacht Armee Mann Wellington]	6691	26%	1	+7%	893	+1%	1119	+1%
3	T71: [Deutschland Krieg Preußen Volk Napoleon Frankreich Macht Frieden Europa Land]	5236	20%	0	-21%	525	-5%**	778	-2%
4	T35: [König Bismarck Wilhelm Kaiser General Minister Stein Berlin Graf Moltke]	3669	14%	2	+52%	324	-5%**	529	-1%
5	T156: [Schlacht Sieg Feind Heer König Mann Kampf Tag Tapferkeit Franzose]	3517	13%	0	-14%	506	+2%	546	+0%
6	T176: [Frankreich England Rußland Deutschland Preußen Krieg Italien Spanien Schweden Holland]	3164	12%	0	-13%	506	+3%*	554	+1%
7	T54: [Staat Zeit Volk Deutschland Leben Reich Jahrhundert Macht Entwicklung Gebiet]	3070	12%	0	-12%	219	-5%**	366	-3%**
8	T59: [Tod Leben Volk Herz Freund Mann Wort König Tag Feind]	2903	11%	0	-11%	344	-1%	410	-1%
9	T81: [Herz Himmel Gott Welt Lied Leben Auge Erde Land Nacht]	2718	10%	0	-11%	355	+0%	382	-1%
10	T67: [Preußen Bund Staat König Regierung Deutschland Verfassung Frankfurt Reichstag Bundestag]	2645	10%	2	+56%	371	+1%	412	+0%
11	T140: [Stadt Franzose Feind Festung Truppe Tag Mann Paris Belagerung Angriff]	2602	10%	0	-10%	395	+2%	454	+1%

Exportieren von Analyseergebnissen



Importieren des CSV-Formats
(comma-separated-values)
Beispiel LibreOffice Calc.



Diagramm: Exportieren der Werte im Zeitverlauf

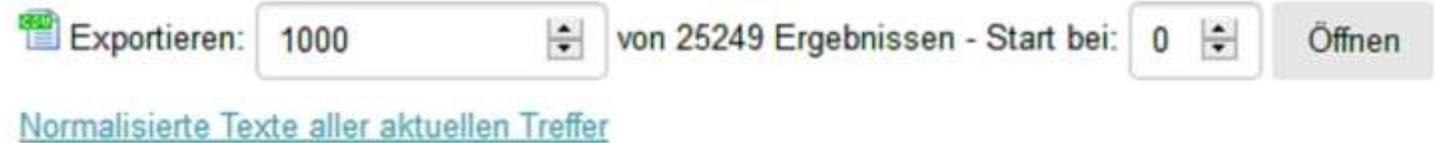
26	Brandenburg-Preussen	175	0%	0	+0%
27	Reformation	15	0%	0	+0%
28	Weltkrieg	2592	1%	0	-1%
29	Gesellschaftskunde	2491	1%	0	-1%

Werte als CSV exportieren

Vergleichsmatrix: Exportieren der Tabelle (ohne Vergleich +/-)

Achtung: Dateiname und Dateiendung müssen selbst vergeben werden (*.csv)

Exportieren von Dokumenten



Formular:

Export beliebiger Anzahl der Treffer (nach aktueller Sortierung, öffnet in neuem Fenster)

Link: Normalisierte Texte aller aktuellen Treffer

Direkt speichern, Alle Treffer, Text normalisiert wie für Topic Modeling verwendet

Weitere Hinweise in PDF: Export von Treffern als CSV zur Analyse in externen Werkzeugen

mit Beispielanalyse in ConText

LibreOffice-Macro, mit dem die CSV-Dateien in einzelne Textdateien geteilt werden können

Dokumentation von Ergebnissen

Aktuelle Seite als Lesezeichen speichern – Titel anpassen

Screenshots (Alt-Druck)

Daten aus Tabellen kopieren

Datentabelle (Legende) in Diagramm

Vergleich Kategorie x Kategorie: Alles markieren (Strg-A) und kopieren

Hochauflösende Diagramme

Als SVG speichern: Crowbar2 (!) <http://nytimes.github.io/svg-crowbar/>

Export der Textseiten im zweiten Teil des Tutorials

Institutionen & Mitarbeiter

Geisteswissenschaften

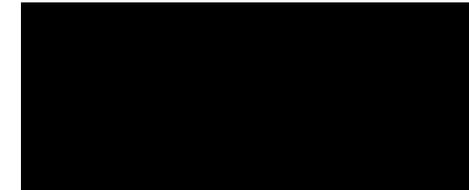
**GEORG ECKERT
INSTITUT**

Leibniz-Institut für internationale
Schulbuchforschung

Prof. Dr. Simone Lässig
Prof. Dr. Ernesto William De Luca
Andreas Weiß
Maik Fiedler

Robert Strötgen (assoziiert)

Computerlinguistik/Informatik



Prof. Dr. Irina Gurevych
Carsten Schnober

Dr. Richard Eckart de Castilho (assoziiert)

Informationswissenschaft

Prof. Dr. Christa Womser-Hacker
Prof. Dr. Thomas Mandl
Dr. Ben Heuwing

