

Bibliometric Review – eine Anleitung zur Erschließung von Forschungsthemen

Arbeitsschritt 1 von 3: Datenrecherche und Datenerhebung

Zitiervorschlag:

Schneiderberg, Christian und Steinhardt, Isabel (2020). Bibliometric Review – eine Anleitung zur systematischen Erschließung von Forschungsthemen. Arbeitsschritt 1: Datenrecherche und Datenerhebung. Kassel. <https://doi.org/10.5281/zenodo.3919371>

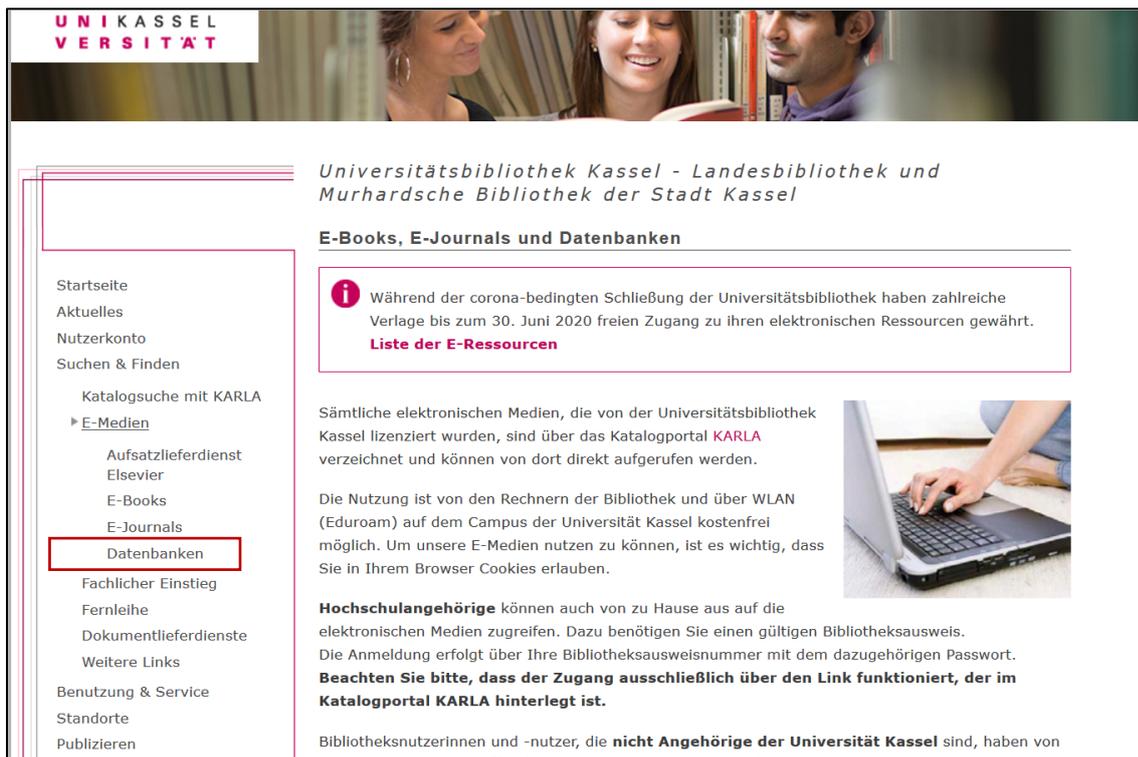
Das Erschließen von Forschungsthemen, also welche Literatur bei einem Thema besonders häufig zitiert wurde oder wer grundsätzlich in einem Forschungsthema publiziert, ist besonders bei Abschlussarbeiten relevant. Zur Beantwortung dieser Forschungsfragen eignen sich bibliometrische Verfahren wie die *Co-Citation Analysis*, das *Bibliographic Coupling* oder die *Co-Occurrence Analyse* von Wörtern. Zur Visualisierung der Analyse kann dann das Open Source Programm [VosViewer](#) genutzt werden, das eine *Map* erstellt, die Zusammenhänge im Forschungsthema durch *Cluster* darstellt.

Zur systematischen Anleitung der Erschließung von Forschungsthemen haben wir eine Anleitung geschrieben, die sich in drei Arbeitsschritte unterteilt. Vorliegend geben wir einen Überblick über die Datenrecherche und Datenerhebung. Nachdem die Daten erhoben wurden bitte mit der zweiten Anleitung zur Datenbereinigung und dann mit der dritten Anleitung zur Datenauswertung weitermachen.

Schritt 1.1

Für die Datenerhebung nutzen wir Web of Science (WoS). WoS ist ein kommerzieller Anbieter, der aber über die meisten Hochschulbibliotheken (meist bei Datenbanken) genutzt werden kann. Wenn Ihr den Zugang nicht direkt findet, dann fragt bei Eurer Bibliothek nach. Achtet darauf, dass Ihr über einen VPN-Client als Mitglied Eurer Hochschule identifizierbar seid.

Wir zeigen hier, wie an der Universität Kassel (da wir beide Lehrende dieser Universität sind) WoS erreicht werden kann. Hierzu über die [Universitätsbibliothek](#) den Reiter „Datenbanken“ auswählen.



**UNI KASSEL
VERSITÄT**

*Universitätsbibliothek Kassel - Landesbibliothek und
Murhardsche Bibliothek der Stadt Kassel*

E-Books, E-Journals und Datenbanken

i Während der corona-bedingten Schließung der Universitätsbibliothek haben zahlreiche Verlage bis zum 30. Juni 2020 freien Zugang zu ihren elektronischen Ressourcen gewährt.
Liste der E-Ressourcen

Sämtliche elektronischen Medien, die von der Universitätsbibliothek Kassel lizenziert wurden, sind über das Katalogportal **KARLA** verzeichnet und können von dort direkt aufgerufen werden.

Die Nutzung ist von den Rechnern der Bibliothek und über WLAN (Eduroam) auf dem Campus der Universität Kassel kostenfrei möglich. Um unsere E-Medien nutzen zu können, ist es wichtig, dass Sie in Ihrem Browser Cookies erlauben.

Hochschulangehörige können auch von zu Hause aus auf die elektronischen Medien zugreifen. Dazu benötigen Sie einen gültigen Bibliotheksausweis. Die Anmeldung erfolgt über Ihre Bibliotheksausweisnummer mit dem dazugehörigen Passwort. **Beachten Sie bitte, dass der Zugang ausschließlich über den Link funktioniert, der im Katalogportal KARLA hinterlegt ist.**

Bibliotheksnutzerinnen und -nutzer, die **nicht Angehörige der Universität Kassel** sind, haben von

Startseite
Aktuelles
Nutzerkonto
Suchen & Finden
Katalogsuche mit KARLA
► **E-Medien**
Aufsatzlieferdienst
Elsevier
E-Books
E-Journals
Datenbanken
Fachlicher Einstieg
Fernleihe
Dokumentlieferdienste
Weitere Links
Benutzung & Service
Standorte
Publizieren



Schritt 1.2

Als nächsten Schritt bei "Schnelle Suche" Web of Science eingeben.

The screenshot shows the 'Datenbank-Infosystem (DBIS)' search page. The search bar under 'Schnelle Suche' contains the text 'web of science' and is circled in red. The search criteria are: 'Suche über alle Felder: "web of science"'. The search results table is as follows:

Fachgebiete	Anzahl
Allgemein / Fachübergreifend	1231
Allgemeine und vergleichende Sprach- und Literaturwissenschaft	322
Anglistik, Amerikanistik	202
Archäologie	171
Architektur, Bauingenieur- und Vermessungswesen	241
Biologie	339
Chemie	177
Elektrotechnik, Mess- und Regelungstechnik	68
Energie, Umweltschutz, Kerntechnik	166
Ethnologie (Volks- und Völkerkunde)	173
Geographie	252
Geowissenschaften	154
Germanistik, Niederländische Philologie, Skandinavistik	487
Geschichte	1551
Informatik	83
Informations-, Buch- und Bibliothekswesen, Handschriftenkunde	256
Klassische Philologie	113
Kunstgeschichte	393
Land- und Forstwirtschaft, Gartenbau, Fischereiwirtschaft, Hauswirtschaft, Ernährung	216
Maschinenwesen, Werkstoffwissenschaften, Fertigungstechnik, Bergbau und Hüttenwesen, Verkehrstechnik, Feinwerktechnik	111
Mathematik	66
Medien- und Kommunikationswissenschaften, Publizistik, Film- und Theaterwissenschaft	288

Schritt 1.3

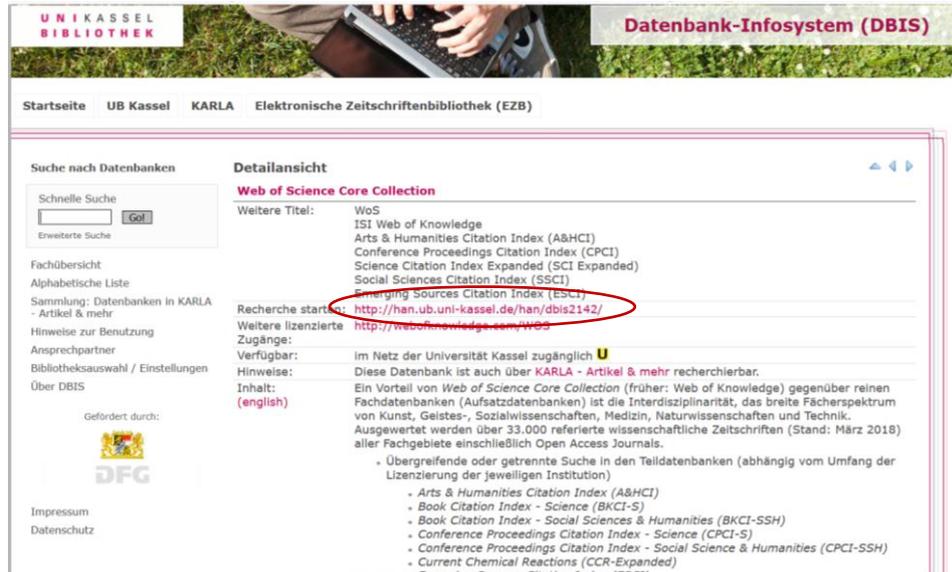
Als nächstes "Web of Science core collection" auswählen.

The screenshot shows the search results page for 'web of science'. The search criteria are: 'Suche über alle Felder: "web of science"'. The search results table is as follows:

Gewählte Datenbanken (10 Treffer)	Zugang
Highly cited researchers	frei im Web F
ip12: Information You Can Trust	frei im Web F
MEDLINE	Uninetz U
Neuroscience Information Framework	frei im Web F
PubMed	frei im Web F
SocioSite	frei im Web F
Web of Science / Arts & Humanities Citation Index	Uninetz U
Web of Science Core Collection	Uninetz U
Web of Science / Science Citation Index Expanded	Uninetz U
Web of Science / Social Sciences Citation Index	Uninetz U

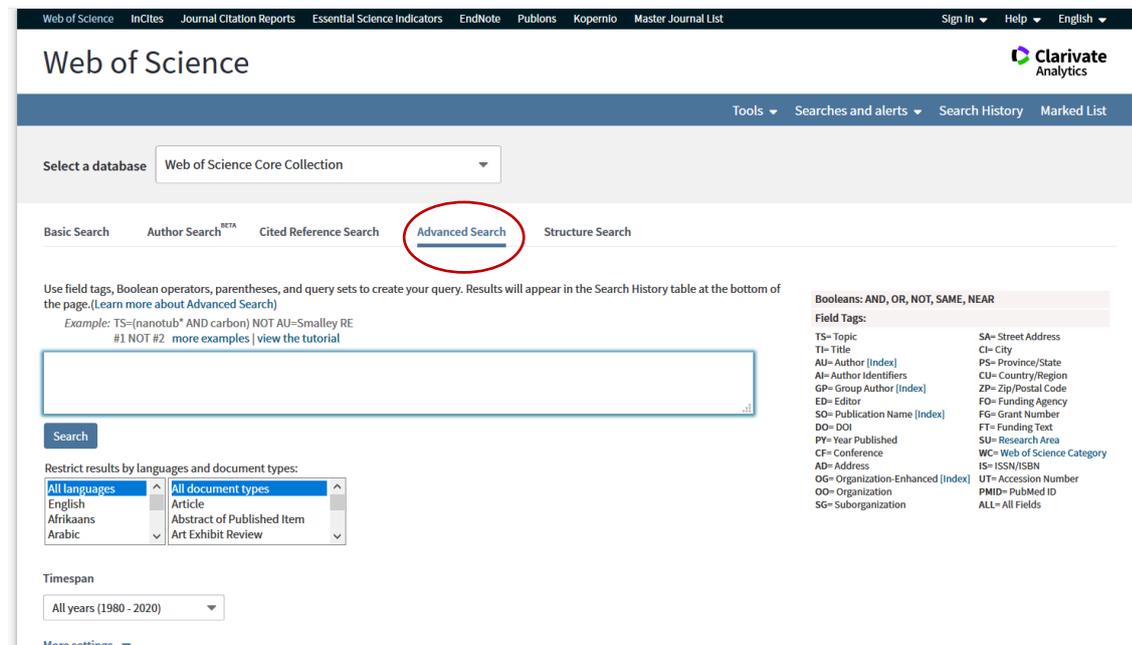
Schritt 1.4

Nun müsst Ihr auswählen, dass ihr die „Recherche starten“ wollt, dazu auf den Link gehen, der automatisch den Zugang zum Web of Science über die UB Uni Kassel regelt. Die aufgezogenen Schritte sind wahrscheinlich je nachdem an welcher Hochschule ihr seid unterschiedlich. Setzt Euch dafür am besten mit Eurer Bibliothek in Verbindung.



Schritt 1.5

Bei WoS könnt Ihr nun die Suche nach Artikeln starten. Da wir vorher „Core Collection“ ausgewählt haben, suchen wir in der Datenbank, die alle Spezialdatenbanken von WoS integriert. Als nächsten Schritt nun „Advanced search“ anklicken.



Schritt 1.6

Als nächstes kann nun nach Artikeln gesucht werden. Durch die *Advanced Search* gibt es eine ganze Reihe an Möglichkeiten (siehe roter Kasten), z. B. nach Autoren, Zeitschriften oder Konferenzen zu suchen. Für unser Vorhaben ein Forschungsthema zu untersuchen eignen sich am besten zwei Möglichkeiten:

1. Eine Schlagwortsuche entweder im Titel (=TI im Kasten) oder nach Topics (=TS im Kasten)
2. Die andere Möglichkeit ist sich ganze Zeitschriften anzuschauen, von denen man weiß, dass sie für ein Forschungsthema zentral sind.

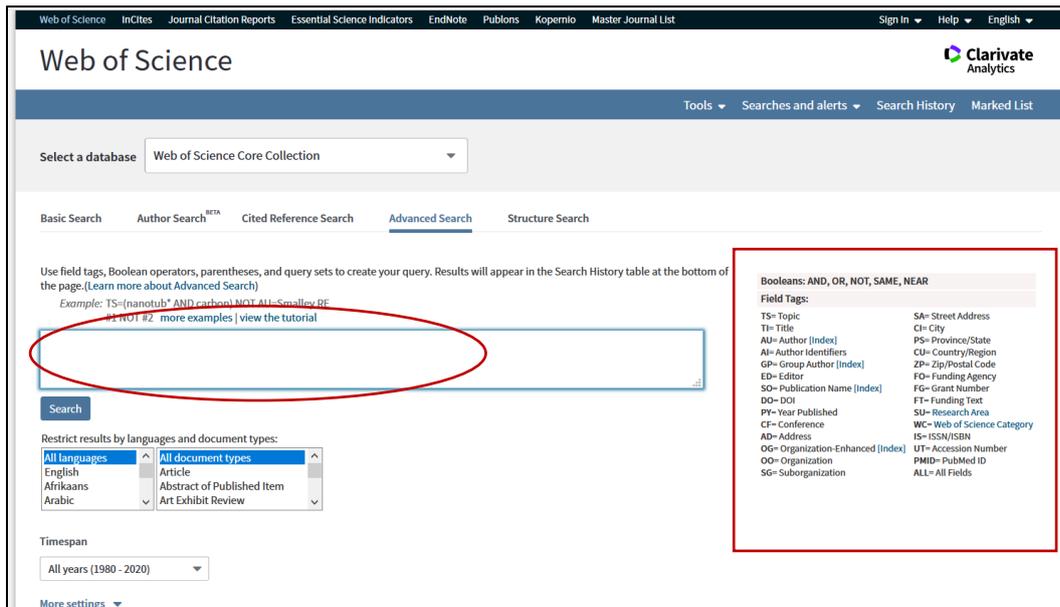
Im Folgenden werden wir Variante 1 beschreiben. Hierzu nutzen wir die Suche via Topics (TS). Die Topic-Suche hat den Vorteil, dass im Titel, im Abstract und in den Keywords (Schlagworten) des Artikels nach unseren Suchwörtern gesucht wird. Bei der Topic-Suche können Wörter miteinander kombiniert werden über AND. Zudem kann man mehrere Suchwortkombinationen durch OR miteinander verbinden und muss dadurch nicht mehrere Suchdurchläufe starten. Zudem können Suchwörter auch durch den Befehl NOT ausgeschlossen werden. Zudem können der Singular und Plural von Wörtern durch * (Sternchen) inkludiert werden.

Für die Anleitung haben wir das Forschungsthema digitalen Textanalyse gewählt, an dem wir gerade am Arbeiten sind. Um Forschungsthemen zu analysieren ist es zentral zu wissen, worum es in diesen geht. Denn die Suchwörter müssen den Kern des Forschungsthemas treffen, sonst habt Ihr viele Treffer, die für Eure Untersuchung nicht relevant sind (*Noise*). Oder in Euren Daten fehlen Texte, die aber eigentlich relevant wären. Eine Möglichkeit besteht darin sich bereits gefundene Artikel Eures Forschungsthemas anzusehen und sich die zentralen Begrifflichkeiten aus Titel und Abstract zu notieren. Wichtig ist zudem bei den ersten Trefferlisten in WoS zu schauen, ob die Treffer relevant für Eurer zu untersuchendes Forschungsthema sind oder nicht. **Ihr seht: es ist wichtig das Feld schon zu kennen, sonst könnt Ihr keine sinnvollen Entscheidungen treffen!**

In unserem Beispiel haben wir folgende Suchwortkombination genutzt:

```
TS=(„digital*“ AND „content analys*“ OR „analys* of content*“) OR TS=(„digital*“ AND „text* analys*“ OR „analys* of text*“) OR TS=(„digital*“ AND „document analys*“ OR „analys* of document*“) OR TS=(„machine learn*“ AND „content analys*“ OR „analys* of content*“) OR TS=(„machine learn*“ AND „text analys*“ OR „analys* of text*“) OR TS=(„machine learn*“ AND „document analys*“ OR „analys* of content*“) OR TS=(„blended reading“ OR „text mining“ OR „distant reading“ OR „culturomics“ OR „topic modeling“)
```

Die Suchwortkombination kann nun in das rot umkreiste Feld eingegeben werden. ACHTUNG: Es ist wichtig auf die „ (Anführungszeichen) zu achten, ebenso auf die Klammern, sonst kann WoS die Suche nicht korrekt ausführen und Ihr bekommt eine Fehlermeldung.

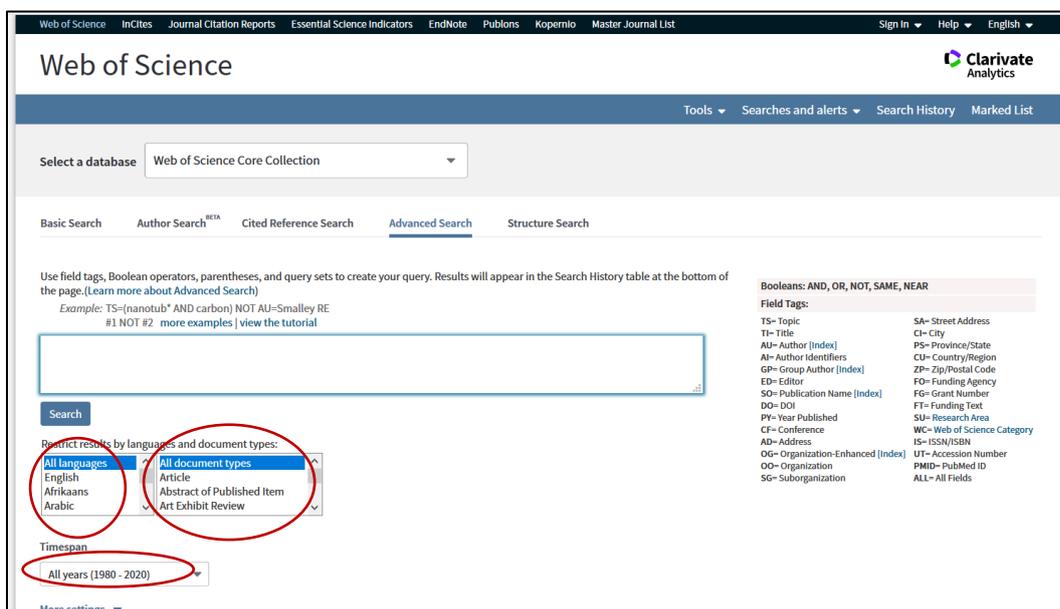


Schritt 1.7

Nachdem ihr die Suchwortkombination eingegeben habt, könnt Ihr als nächsten Schritt weitere Einstellungen vornehmen:

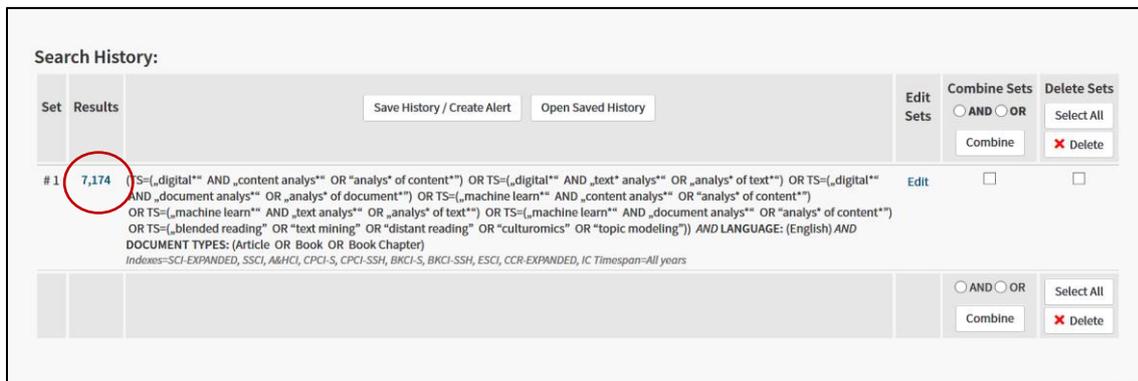
1. Sprache wählen: Das *Web of Science* ist großmehheitlich eine Datenbank englischsprachiger Beiträge. Entsprechend ist es sinnvoll die Sprache Englisch auszuwählen, damit ein umfassenderes Ergebnis der vorhandenen Literatur. Sollen mehrere Sprachen („*language*“) „Strg-Taste“ drücken und anklicken mit Maus.
2. Dokumententyp („*document type*“) wählen: Bei Auswahl mehrerer Dokumententypen: „Strg-Taste“ drücken und anklicken mit Maus. Für die Literaturrecherche in den Sozial- und Geisteswissenschaften ist die Einschränkung der Suche auf folgende Dokumententypen empfehlenswert: Zeitschriftenbeiträge („*Article*“), Bücher („*Book*“), Beiträge in Sammelwerken („*Book chapter*“)
3. Zeitraum wählen: Wenn Ihr beispielsweise eine Analyse der letzten zehn Jahre vornehmen wollt, dann könnt Ihr diese Zeitspanne hier auswählen.

Nachdem Ihr die Auswahl getroffen habt auf „**search**“ klicken.

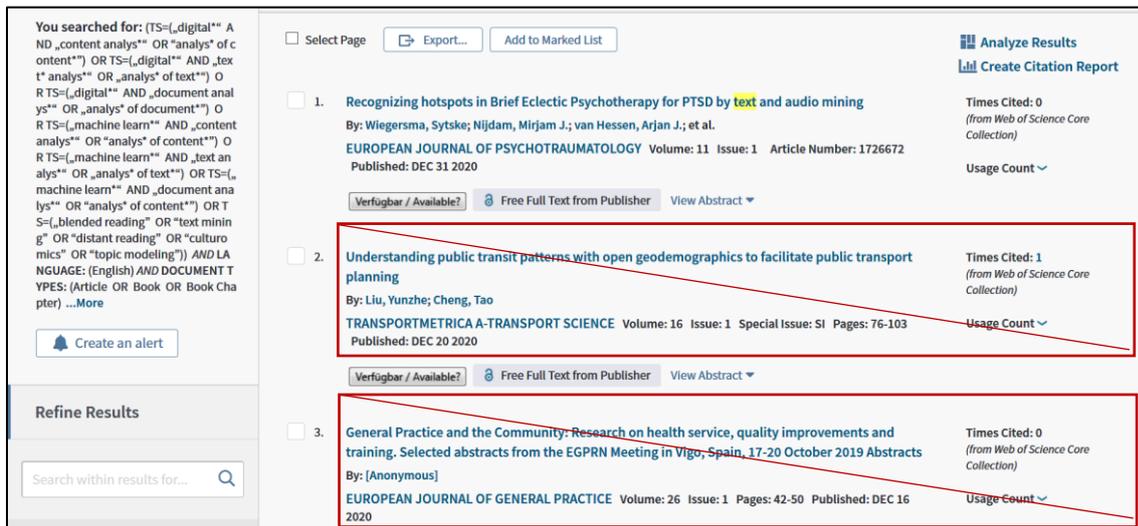


Schritt 1.8

Unter der Suchmaske erscheint nun die „Search History“. In Spalte **“Results“** seht Ihr die Anzahl der Treffer. Wenn Ihr auf die Zahl klickt wird Euch eine Liste der Treffer angezeigt.



In dem vorliegenden Beispiel haben wir über die Trefferliste schnell festgestellt, dass in dem Datensatz noch zu viele Treffer sind, die nicht zu dem Forschungsthema gehören (siehe rote Kästen). Deshalb musste die Suche spezifiziert werden.



Schritt 1.9

Für das vorliegende Beispiel gab es zwei Möglichkeiten:

1. Option: Veränderung der Schlagwortsuche
2. Option: Suche einschränken auf Publikationstitel.

Da es sich bei unserem Forschungsthema um methodische Fragen und die Anwendung von Methoden drehen soll, haben wir uns für die zweite Option entschieden, also nur noch im Titel zu suchen. Hintergrund ist die Annahme, dass Publikationen über Methoden bzw. wenn die Anwendung einer Methode für die Publikation zentral ist, dass die Methode dann auch im Titel der Publikation erscheint.

Die neue Suchwortkombination sieht entsprechend so aus:

TI=(„digital“ AND „content analys“ OR „analys* of content“) OR TI=(„digital“ AND „text* analys“ OR „analys* of text“) OR TI=(„digital“ AND „document analys“ OR „analys* of document“) OR TI=(„machine learn“ AND „content analys“ OR „analys* of content“) OR TI=(„machine learn“ AND „text analys“ OR „analys* of text“) OR TI=(„machine learn“ AND „document analys“ OR „analys* of content“) OR TI=(„blended reading“ OR „text mining“ OR „distant reading“ OR „culturomics“ OR „topic modeling“) OR TI=(„digital“ AND „Inhaltsanalyse“) OR TI=(„digital“ AND „Textanalyse“) OR TI=(„digital“ AND „Dokumentenanalyse“) OR TI=(„machine learn“ AND „Inhaltsanalyse“) OR TI=(„machine learn“ AND „Textanalyse“) OR TI=(„machine learn“ AND „Dokumentenanalyse“)

Die modifizierte Suche ergab 1.517 Treffer.

Search History:

Set	Results		Edit Sets	Combine Sets	Delete Sets
		Save History / Create Alert	Open Saved History	<input type="radio"/> AND <input type="radio"/> OR Combine	Select All Delete
# 2	1,517	TI=(„digital“ AND „content analys“ OR „analys* of content“) OR TI=(„digital“ AND „text* analys“ OR „analys* of text“) OR TI=(„digital“ AND „document analys“ OR „analys* of document“) OR TI=(„machine learn“ AND „content analys“ OR „analys* of content“) OR TI=(„machine learn“ AND „text analys“ OR „analys* of text“) OR TI=(„machine learn“ AND „document analys“ OR „analys* of content“) OR TI=(„blended reading“ OR „text mining“ OR „distant reading“ OR „culturomics“ OR „topic modeling“) OR TI=(„digital“ AND „Inhaltsanalyse“) OR TI=(„digital“ AND „Textanalyse“) OR TI=(„digital“ AND „Dokumentenanalyse“) OR TI=(„machine learn“ AND „Inhaltsanalyse“) OR TI=(„machine learn“ AND „Textanalyse“) OR TI=(„machine learn“ AND „Dokumentenanalyse“) AND LANGUAGE: (English OR German) AND DOCUMENT TYPES: (Article OR Book OR Book Chapter) <small>Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years</small>	Edit	<input type="checkbox"/>	<input type="checkbox"/>
# 1	7,174	TS=(„digital“ AND „content analys“ OR „analys* of content“) OR TS=(„digital“ AND „text* analys“ OR „analys* of text“) OR TS=(„digital“ AND „document analys“ OR „analys* of document“) OR TS=(„machine learn“ AND „content analys“ OR „analys* of content“) OR TS=(„machine learn“ AND „text analys“ OR „analys* of text“) OR TS=(„machine learn“ AND „document analys“ OR „analys* of content“) OR TS=(„blended reading“ OR „text mining“ OR „distant reading“ OR „culturomics“ OR „topic modeling“) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book OR Book Chapter) <small>Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years</small>	Edit	<input type="checkbox"/>	<input type="checkbox"/>
				<input type="radio"/> AND <input type="radio"/> OR Combine	Select All Delete

Schritt 1.10

Bei der erneuten Durchsicht der Trefferliste zeigt sich, dass die gefundenen Publikationen relevant sind, weshalb mit diesem Korpus (=enthaltene Angaben zu den Publikationen) gearbeitet wird. Dazu muss nun der Korpus heruntergeladen werden. Dazu auf „Export“ klicken und dann „Other file formats“ auswählen.

The screenshot shows the Web of Science search results page. The search results are sorted by Date. The 'Export...' button is highlighted with a red circle. A dropdown menu is open, showing various export options, with 'Other File Formats' also highlighted by a red circle. The page displays search results for '1,517' records, including a list of articles with their titles, authors, and publication details.

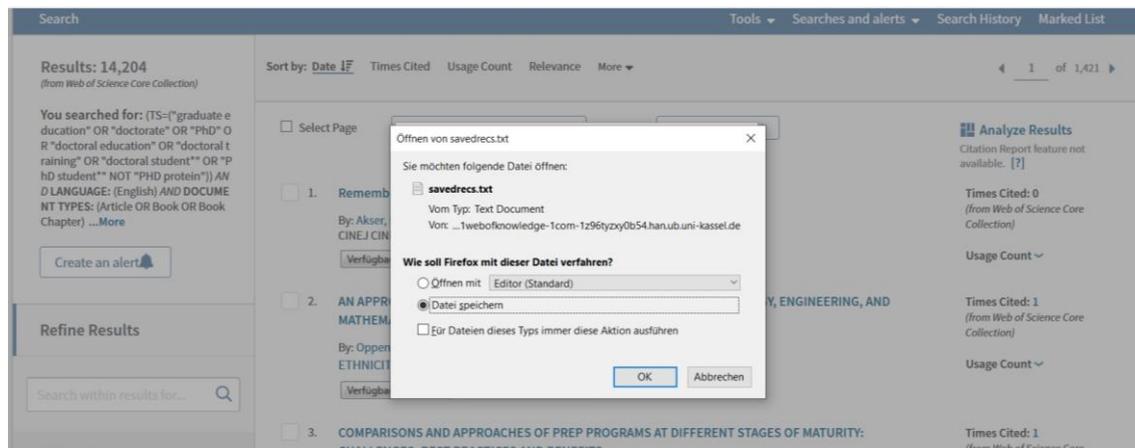
Schritt 1.11

In dem aufscheinenden Fenster dann „Records from ... to ...“ auswählen. Dies hat den Hintergrund, dass z.B. die Universität Kassel nur eine Lizenz hat 500 Records (also Publikationsangaben) auf einmal herunterzuladen. Deshalb muss dieser Vorgang mehrmals wiederholt werden. Also dann die Treffer 501-1000, 1001-1500 usw. heruntergeladen werden, bis der gesamte Korpus heruntergeladen wurde. Bei „Record Content“ bitte „Full record and cited references“ auswählen. Bei „File format“ bitte „tab-delimited (WIN, UTF-8)“ auswählen und dann auf „export“ klicken.

The screenshot shows the 'Export Records to File' dialog box in the Web of Science interface. The dialog box has three main sections: 'Records from: 1 to 500' (with a note 'No more than 500 records at a time.'), 'Record Content: Full Record and Cited References', and 'File Format: Tab-delimited (Win, UTF-8)'. The 'Export' button is visible at the bottom right of the dialog box.

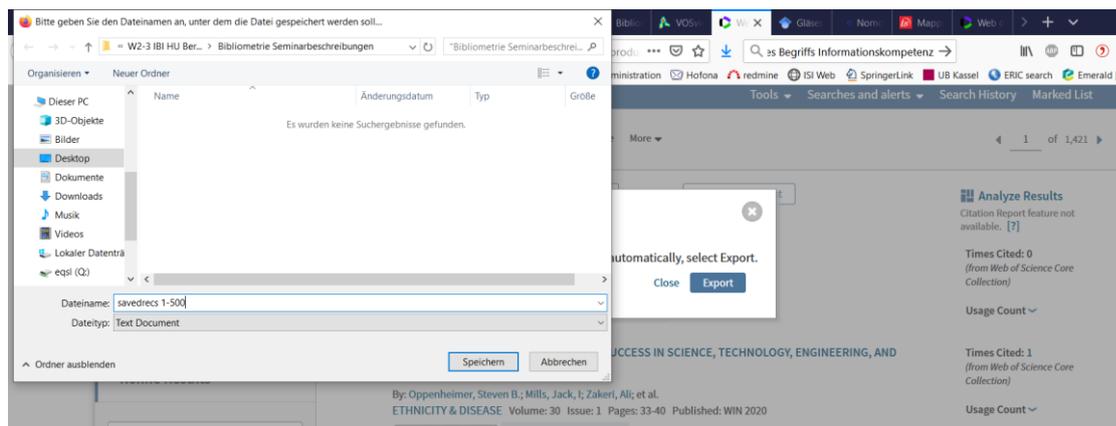
Schritt 1.12

Im nächsten Fenster **“Datei speichern“** auswählen.



Schritt 1.13

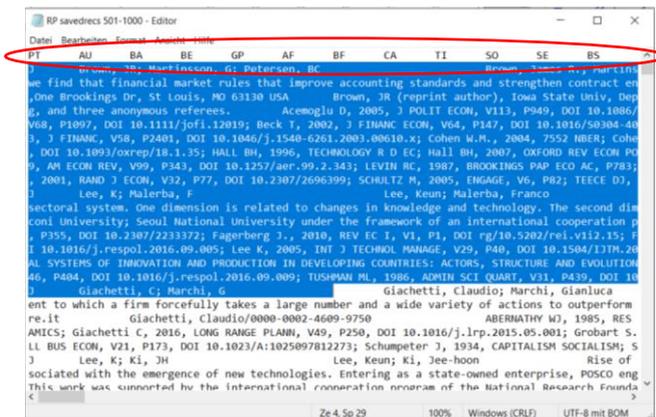
Dann einen Dateiordner auswählen und die Datei benennen. Wichtig ist dabei zu markieren um welchen Teil des Datenexportes es sich handelt, also am besten 1-500, 501-1000 usw. an den Dateinamen anfügen. Die Schritte 1.10-1.13 sooft wiederholen, bis der gesamte Korpus heruntergeladen wurde.



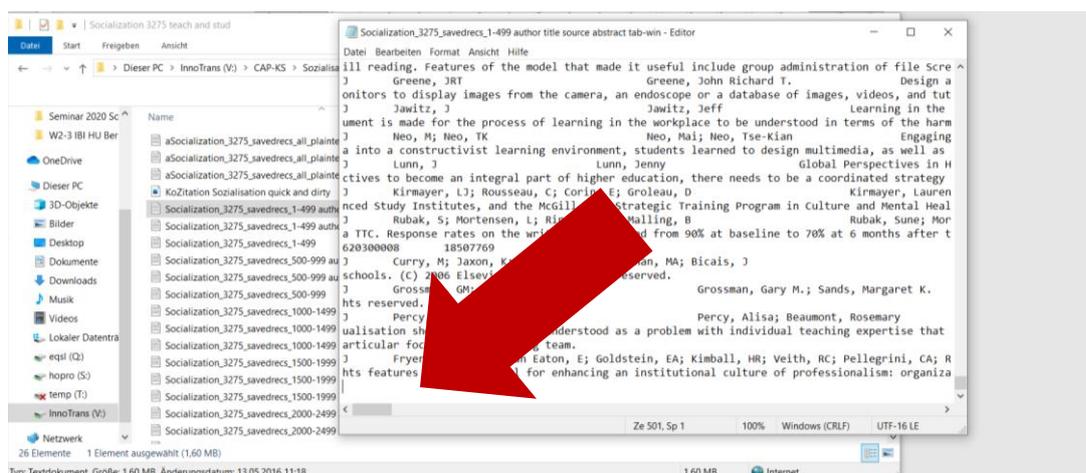
Schritt 1.14

Die heruntergeladenen Daten müssen nun in einer Datei zusammengefügt werden.

1. Die Datei mit den Daten 1-500 öffnen (Doppelklick auf die Datei, so dass sich der Editor öffnet),
2. die Datei unter einem neuen Namen speichern (z. B. „Name der Datei“ gesamt),
3. ans Ende der Daten scrollen;
4. die Datei mit den Daten 501-1000 öffnen,
5. außer Überschrift alles markieren (z. B. Maus-Cursor nach unten ziehen oder strg + a),



6. alles kopieren (strg + c)
7. und in letzter Zeile bei txt-Datei „gesamt“ einfügen (strg + v)



8. Schritte 3-7 sooft wiederholen, bis alle Daten in einer txt-Datei vereinigt sind.

Exkurs: Sofortanalyse (unbereinigte Daten!)

Bevor Ihr die Daten herunterladet kann der Korpus mit den Analysetools von WoS einer ersten Sofortanalyse unterzogen werden. **ACHTUNG:** Es handelt sich dabei um unbereinigte Daten, d. h., es können im Datensatz noch Fehler sein. Für die Bereinigung des Datensatzes siehe die Anleitung „Arbeitsschritt 2 von 3 Datenbereinigung“.

Um eine Sofortanalyse durchzuführen auf „**Analyze Results**“ klicken.

The screenshot displays the Web of Science interface. At the top right, the Clarivate Analytics logo is visible. The main header includes 'Web of Science' and navigation options like 'Tools', 'Searches and alerts', 'Search History', and 'Marked List'. The search results section shows 14,204 results from the Web of Science Core Collection. The search criteria are: (TS=("graduate education" OR "doctorate" OR "PHD" OR "doctoral education" OR "doctoral training" OR "doctoral student" OR "PHD student" NOT "PHD protein")) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article OR Book OR Book Chapter) ...More. A 'Create an alert' button is present. The results are sorted by Date (1 of 1,421). The first three results are listed, each with a 'Verfügbar / Available?' button and a 'Free Full Text from Publisher' button. The 'Analyze Results' button is circled in red, with a tooltip indicating 'Citation Report feature not available. [?]'. The 'Refine Results' section at the bottom left has a search box for refining results.

Mit der Sofortanalyse können einfache Metriken wie zu Publikationsjahr, Autor*innen oder Länder erstellt werden. Über **download „all data rows“** können alle Daten auch heruntergeladen werden.

The screenshot shows the 'Web of Science' interface with a search results analysis for 'Publication Years'. The sidebar on the left contains various filters, with 'Publication Years', 'Authors', and 'Countries/Regions' highlighted by red circles. The main area displays a bar chart and a table of results. The table has columns for 'Field: Publication Years', 'Record Count', '% of 1,517', and 'Bar Chart'. The 'Download' button for 'All data rows (up to 100,000)' is highlighted with a red circle.

Select	Field: Publication Years	Record Count	% of 1,517	Bar Chart
<input type="checkbox"/>	2020	78	5.142 %	■
<input type="checkbox"/>	2019	275	18.128 %	■
<input type="checkbox"/>	2018	202	13.316 %	■
<input type="checkbox"/>	2017	150	9.918 %	■
<input type="checkbox"/>	2016	151	9.954 %	■
<input type="checkbox"/>	2015	124	8.174 %	■
<input type="checkbox"/>	2014	86	5.669 %	■
<input type="checkbox"/>	2013	75	4.944 %	■
<input type="checkbox"/>	2012	77	5.076 %	■
<input type="checkbox"/>	2011	34	2.241 %	■
<input type="checkbox"/>	2010	30	1.978 %	■
<input type="checkbox"/>	2009	36	2.373 %	■
<input type="checkbox"/>	2008	34	2.241 %	■
<input type="checkbox"/>	2007	17	1.121 %	■
<input type="checkbox"/>	2006	26	1.714 %	■
<input type="checkbox"/>	2005	27	1.780 %	■
<input type="checkbox"/>	2004	28	1.846 %	■
<input type="checkbox"/>	2003	15	0.989 %	■
<input type="checkbox"/>	2002	10	0.659 %	■
<input type="checkbox"/>	2001	4	0.264 %	■
<input type="checkbox"/>	2000	5	0.330 %	■
<input type="checkbox"/>	1999	9	0.593 %	■
<input type="checkbox"/>	1998	4	0.264 %	■
<input type="checkbox"/>	1997	3	0.198 %	■
<input type="checkbox"/>	1996	1	0.066 %	■

At the bottom of the table, there is a section for download options:

- Data rows displayed in table
- All data rows (up to 100,000)

The 'Download' button next to the second option is highlighted with a red circle.