Einsteiger-Tutorial zu Korpusrecherchen mit COSMAS II

In diesem Tutorial erfahren wir, wie man Korpusrecherchen mit Hilfe des Korpusabfragesystems COSMAS II durchführt. Über COSMAS II sind verschiedene Korpora durchsuchbar, deren wichtigstes wohl das Deutsche Referenzkorpus (DeReKo) ist, das wiederum aus unterschiedlichen Subkorpora besteht. Es lohnt sich, sich mit der ausführlichen Dokumentation der Korpora vertraut zu machen.
Angenommen, wir wollen im DeReKo den Gebrauch des Erstglieds (Präfixoids?) *Schwarz-* bei Personenbezeichnungen auf *-er* nachvollziehen, z.B. *Schwarzfahrer*, *Schwarzseher*.
Hinweis: Bei der Arbeit mit den exportierten Dateien verwendet dieses Tutorial Notepad++. Wie in Tutorial 01 („Grundlegendes“) dargelegt, funktioniert Notepad++ nur unter Windows. Bei anderen Textverarbeitungsprogrammen wie TextWrangler lassen sich jedoch die gleichen Arbeitsschritte durchführen, auch wenn einige Funktionen ggf. geringfügig anders heißen.

1. Erste Schritte

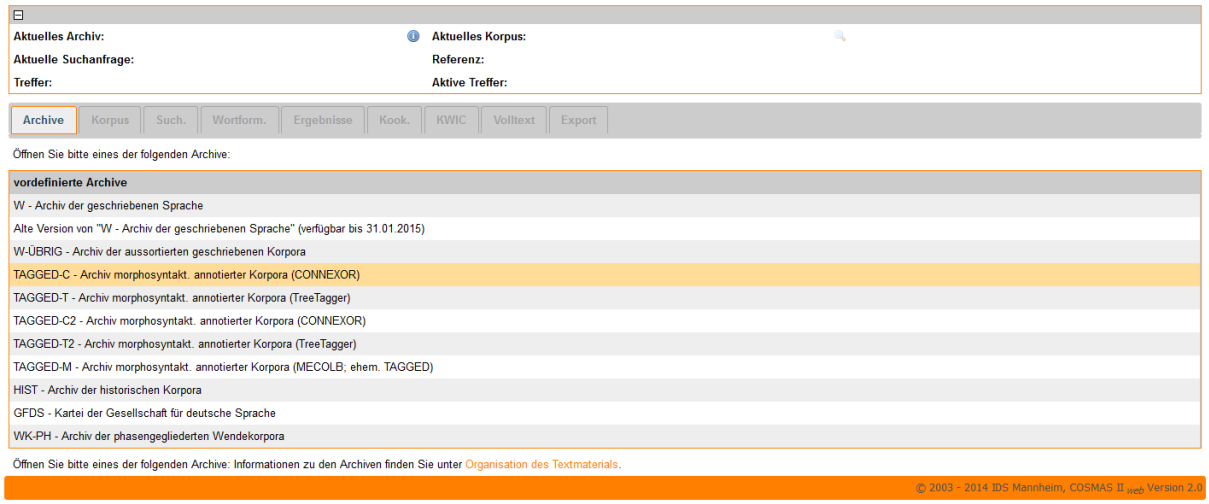
COSMAS II finden Sie unter https://cosmas2.ids-mannheim.de/cosmas2-web/.
Zur Benutzung des Korpus ist einmalig eine **Anmeldung** erforderlich, die jedoch schnell und problemlos möglich ist.

2. COSMAS II$_{win}$ vs. COSMAS II$_{web}$

Sie können COSMAS II entweder online oder mit Hilfe eines Windows-Clients benutzen. Für die meisten Fragestellungen genügt die Online-Variante, auf die sich die folgende Darstellung bezieht. Nur bei einigen komplexeren Suchanfragen ist die Web-Applikation überfordert.

3. Auswählen von Archiv und Korpus

Bei der Suche müssen Sie zunächst ein **Archiv** auswählen. Ein Archiv ist eine Sammlung mehrerer Korpora (und kann ggf. auch selbst als ein einziges, großes Korpus verwendet werden).



Für einfache Suchen zum Gegenwartsdeutschen können Sie gut mit dem größten Archiv, W, arbeiten. Beachten Sie jedoch, dass dieses Archiv recht verschiedenartige Daten enthält, von Zeitungstexten bis zu Wikipedia-Diskussionen; wenn Sie mit einem ausgewogenen Korpus arbeiten möchten, ist es evtl. besser, sich selbst ein Korpus zusammenzustellen.

Wenn Sie indes auf eine morphologische Annotation angewiesen sind, also z.B. nach bestimmten Wortarten (Nomen, Verb etc.) suchen wollen, dann wählen Sie eines der TAGGED-Archive. Für diachrone Fragestellungen, die über die zweite Hälfte des 20. Jh. hinausgehen, ist HIST Ihre Wahl.
Für unsere Recherche benutzen wir das erste **TAGGED-C**-Archiv. (TAGGED-C2 ist übrigens eine Ergänzung mit neueren Texten. Wenn Sie alle Schritte dieses Tutorials erfolgreich durchlaufen haben, versuchen Sie doch einmal, die gleiche Aufgabe noch einmal selbstständig in TAGGED-C2 zu bewältigen...)
Als nächstes müssen Sie ein **Korpus** auswählen:

Vordefinierte Korpora	
TAGGED-C-öffentlich - alle öffentlichen Korpora des Archivs TAGGED-C	
brz - Braunschweiger Zeitung, September 2005 - Dezember 2009	
bvz - Burgenländische Volkszeitung, Januar 2007 - Dezember 2009	
haz - Hannoversche Allgemeine, August 2007 - Dezember 2009	
hmp - Hamburger Morgenpost, Januar 2006 - Dezember 2009	
mm - Mannheimer Morgen, Januar 2006 - Dezember 2009	
non - Niederösterreichische Nachrichten, Januar 2007 - Dezember 2009	
nun - Nürnberger Nachrichten, Januar 2006 - Dezember 2009	
nuz - Nürnberger Zeitung, Januar 2006 - Dezember 2009	
oon - Oberösterreichische Nachrichten, 1999	
rei - Reden und Interviews, Januar 2002 - Dezember 2006	
rhz - Rhein-Zeitung, Januar 2006 - Dezember 2009	
sgt - St. Galler Tagblatt, April 1997 - Dezember 2001, August 2007 - Dezember 2009	
soz - Die Südostschweiz, Januar 2006 - Dezember 2009	
spk - spektrumdirekt, Januar 1997 - Dezember 2009	
van - Vorarlberger Nachrichten, 1999	
vdi - VDI Nachrichten, Februar 2006 - Dezember 2009	
wpd - Wikipedia	

Informationen zu den Korpora finden Sie unter Organisation des Textmaterials.

Hier können Sie Korpora auswählen, die bereits vordefiniert sind. Alternativ können Sie auch ein **eigenes Korpus** zusammenstellen. In vielen Fällen ist das sinnvoll: Zum Beispiel, wenn Sie ein Phänomen nur für lektorierte Zeitungssprache in Deutschland untersuchen wollen und deshalb österreichische und Schweizer Zeitungen sowie Wikipedia-Belege ausschließen wollen.

Hinweis zu Wikipedia: Wenn Sie diachron und mit einem feinkörnigen Zeitraster arbeiten (z.B. ein Wandelphänomen für die ersten 15 Jahre des neuen Jahrtausends untersuchen), sollten Sie die Wikipedia-Belege **immer** ausschließen, da sie alle auf 2011 datiert sind. Im W-Archiv gibt es außerdem ein Archiv mit **Wikipedia-Diskussionen**, das sich hervorragend zur Untersuchung konzeptionell nähesprachlicher Phänomene eignet, aber bei schriftsprachlich orientierten Untersuchungen grundsätzlich ausgeschlossen werden sollte.

Für unsere Recherche gehen wir „opportunistisch“ vor und schließen alle Belege mit ein, wir wählen also das allererste vordefinierte Korpus „TAGGED-C öffentlich“.

4. Suchanfrage

Nun müssen wir unsere Suchanfrage eingeben. Dafür müssen Sie die Grundlagen der COSMAS II-Syntax kennen. Wie in anderen Korpora kann man auch in COSMAS direkt nach Wörtern (Tokens) suchen oder aber nach Lemmas. Benutzt man die Lemmasuche, dann werden automatisch alle Formen gefunden, die zur jeweiligen Grundform gehören, also z.B. für *sein* auch *bin, ist* usw. In der COSMAS-Syntax muss man dafür den Grundformoperator & benutzen, z.B.

```
&sein
```

Ganz wichtig: Das Verhalten des Grundformoperators variiert je nachdem, welche Voreinstellungen in COSMAS II vorgenommen worden sind. Defaultmäßig sucht er nicht nur einfach die Flexionsformen, die zur jeweiligen Grundform gehören (was eigentlich das ist, was man von einem Lemmaoperator erwarten würde), sondern z.B. auch Komposita wie *Seineufer* etc. Das wollen wir in den meisten Fällen nicht, weshalb es sich empfiehlt, die Optionen „Komposita“, „Sonstige Wortbildungsformen“ und „Spezialfälle“ unter Optionen > Lemmatisierung abzuwählen (Fig. 1).
Im weiteren Verlauf dieses Tutorials werden wir allerdings die Lemmasuche nicht nutzen, sondern nur die einfache Wortsuche – da sich aber in sehr vielen Fällen die Nutzung der Lemmainformationen anbietet, schien mir der kurze Exkurs an dieser Stelle wichtig.

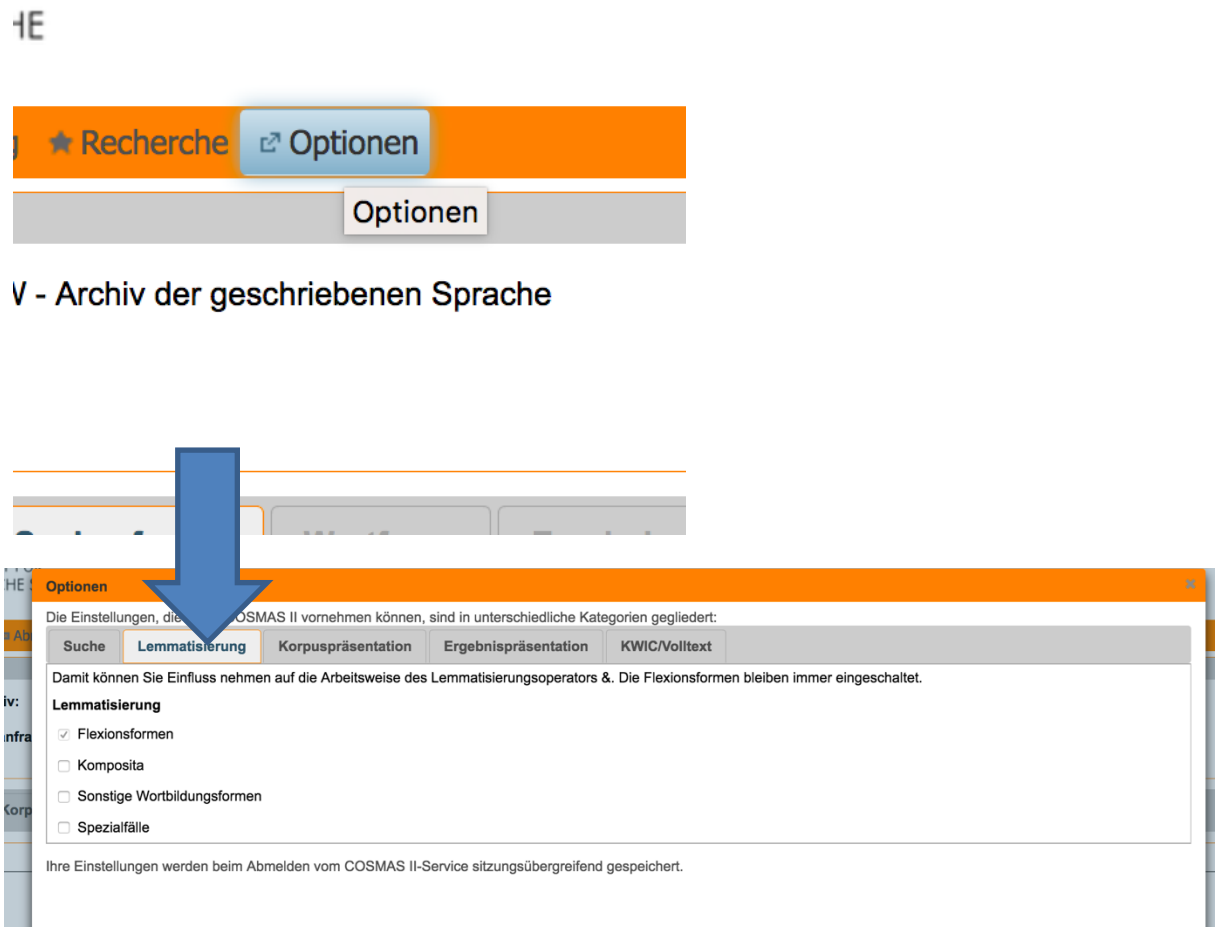


Fig. 1: Einstellen der Funktionsweise des Grundformoperators & in COSMAS II.

Ein weiteres wichtiges Element sind die **Verknüpfungsoperatoren**. Die Wörter „und“ und „oder“ lassen sich als solche benutzen:

```
Tee und Kaffee
```

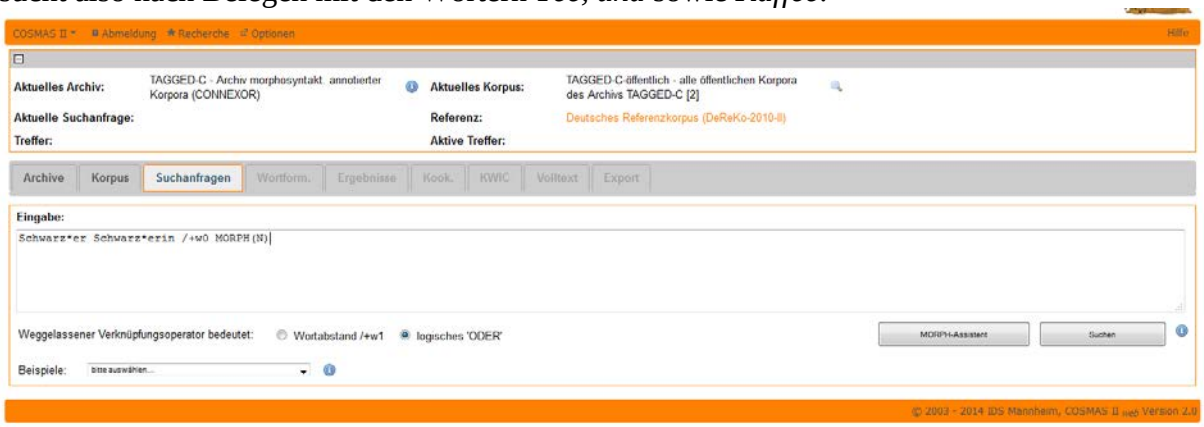
findet alle Belege, in denen das Wort *Tee* und das Wort *Kaffee* vorkommen.

```
Tee oder Kaffee
```

findet Belege, in denen *eines* der beiden Wörter vorkommt. Wichtig: Wenn Sie nach dem **Wort** *und* oder dem **Wort** *oder* suchen wollen, müssen Sie es in der Suchanfrage in Anführungszeichen setzen:

```
Tee "und" Kaffee
```

sucht also nach Belegen mit den Wörtern *Tee, und* sowie *Kaffee*.



Unterhalb des Eingabefeldes können Sie auswählen, wie COSMAS II einen weggelassenen Verknüpfungsoperator (vulgo: ein einfaches Leerzeichen) interpretieren soll. Standardmäßig angewählt ist *logisches ODER*, d.h. „Tee Kaffee“ bringt die gleichen Resultate wie „Tee ODER Kaffee“.
Die Alternative hierzu führt uns zu einem weiteren wichtigen Operator, dem **Wortabstandsoperator**. Mit /+w*n* und /-w*n* können Sie auswählen, in welchem Abstand das Wort nach dem Operator dem Wort vor dem Operator nachfolgen oder vorangehen soll:

```
Tee /+w1 oder /+w1 Kaffee
```

bedeutet, dass *oder* im Abstand von 1 auf *Tee* folgen soll und *Kaffee* im Abstand von 1 auf *oder*. Sie finden so also alle Belege für *Tee oder Kaffee*.

```
Tee /+w3 Kaffee
```

bedeutet, dass *Kaffee* im Abstand von **maximal** 3 Wörtern auf *Tee* folgen soll. Hier finden Sie also Belege für z.B. *Tee oder Kaffee* und *Tee oder einen Kaffee*. Wenn sie das + weglassen, bedeutet dies, dass das Wort sowohl vorangehen als auch nachfolgen kann:

```
Tee /w3 Kaffee
```

findet also Belege für *Tee oder (einen) Kaffee* ebenso wie für *Kaffee oder (einen) Tee*.

Hinweis zu COSMAS II_{win}: Wenn Sie statt mit der Online-Version mit dem Windows-Client arbeiten, müssen Sie die Reihenfolge der Zeichen im Wortabstandsoperator umkehren: Statt `/+w3` geben Sie hier `/+3w` ein. (Das ist sehr irritierend und wird in zukünftigen Versionen hoffentlich vereinheitlicht.)

Auch für unsere Suchanfrage ist der Wortabstandsoperator wichtig, obwohl wir nur ein einziges Wort suchen. Wir wollen nämlich nur Ergebnisse für eine bestimmte Wortklasse finden:

für **Substantive**. Diese werden mit MORPH(N) gesucht. (Komplexere Anfragen können Sie mit dem MORPH-Operator basteln, s. Button unten rechts im obigen Screenshot.) Wenn wir nun aber eingeben

```
Schwarz*er MORPH(N)
```

versteht COSMAS das zweite Element der Suchanfrage als eigenes Wort: Je nachdem, welche Option unten bei „weggelassener Verknüpfungsoperator bedeutet“ eingestellt ist, wird es also entweder interpretiert als

```
Schwarz*er /+w1 MORPH(N),
```

d.h. es werden alle Belege gefunden, in denen auf *Schwarzfahrer, Schwarzarbeiter* usw. ein Substantiv folgt, oder als

```
Schwarz*er oder MORPH(N),
```

d.h. es werden alle Belege gefunden, in denen **entweder** *Schwarzfahrer, Schwarzarbeiter* etc. **oder** irgendein anderes Nomen auftauchen. Deshalb müssen wir COSMAS erklären, dass **das gleiche Wort** diese morphologische Form haben soll. Das gleiche Wort bedeutet: Wortabstand Null. Genau damit ergänzen wir unsere Suchanfrage:

```
Schwarz*er /+w0 MORPH(N),
```

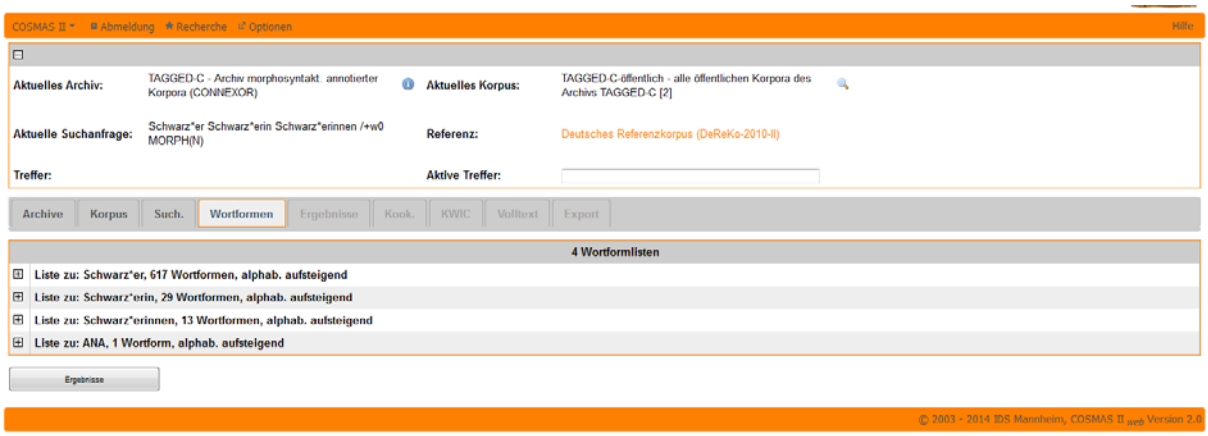
Sie haben sicherlich schon gemerkt, dass * als **Wildcard** fungiert, d.h. für eine beliebige (und beliebig lange) Buchstabenfolge steht. `Schwarz*er` findet also *Schwarzfahrer* ebenso wie *Schwarzseher* oder *Schwarzenegger*.
Um auch die movierte Form zu berücksichtigen, suchen wir auch nach *Schwarzfahrerinnen* und ergänzen unsere Suchanfrage daher:

```
Schwarz*er oder Schwarz*erin oder Schwarz*erinnen /+w1
MORPH(N)
```

Damit ist unsere Suchanfrage vollständig und wir können auf „Suchen“ klicken.

5. Wortformenliste

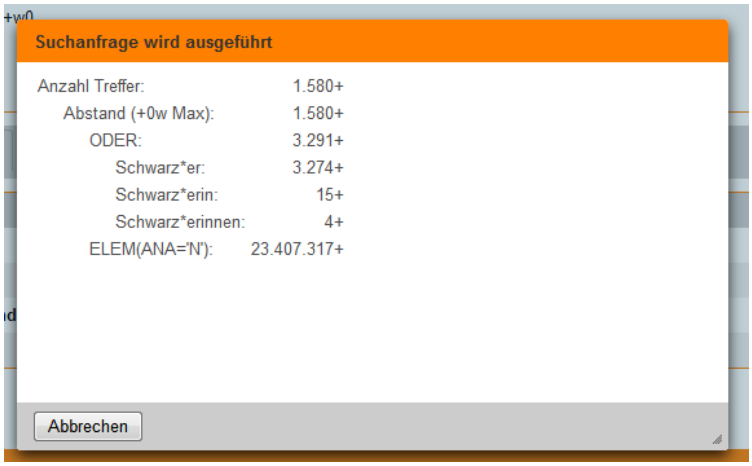
Als erstes Ergebnis erhalten wir die Wortformenliste, d.h. eine Liste aller **Types**, die gefunden wurden. Grundsätzlich könnten wir nun die Types manuell durchgehen und Fehltreffer (z.B. Eigennamen wie *Schwarzenegger*) aus unseren endgültigen Ergebnissen entfernen. Das wäre aber bei insgesamt über 600 Types ziemlich aufwändig, deshalb werden wir das in einem späteren Schritt effizienter erledigen.



Jetzt gehen wir zunächst direkt weiter zu den „Ergebnissen“, indem wir auf den entsprechenden Button klicken.

6. Ergebnisse

Zunächst müssen wir uns etwas gedulden, während die Ergebnisssuche läuft:

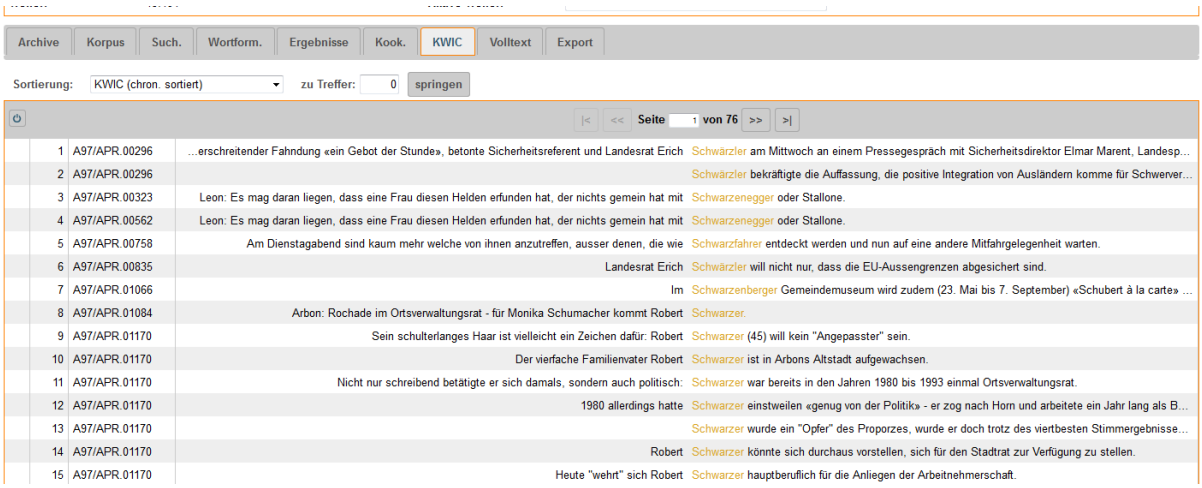


Die Ergebnisse, die wir nun erhalten, sind standardmäßig in der **Quellenansicht** sortiert, d.h. nach der Quelle, der die jeweiligen Belege entstammen.

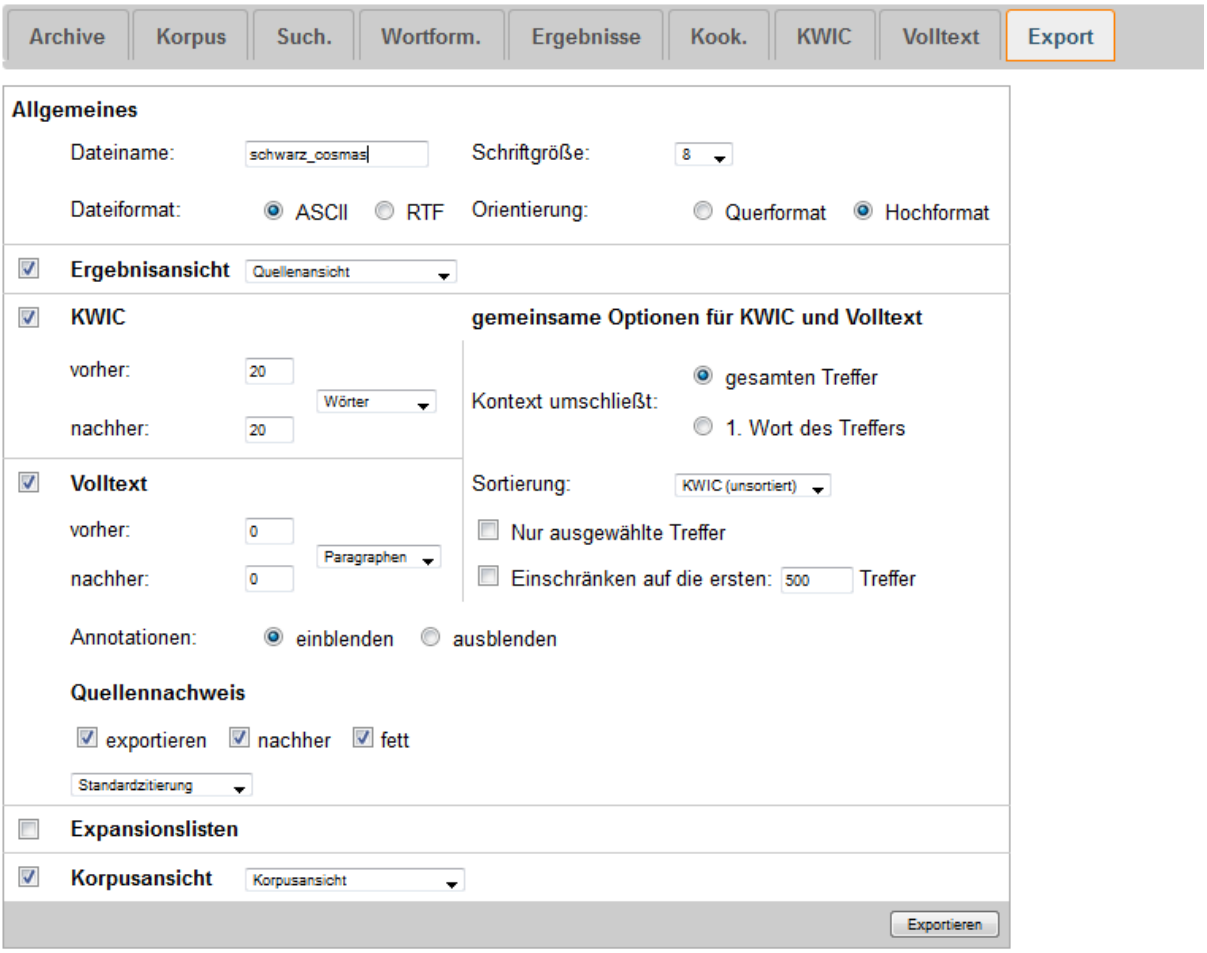
Quellenansicht

Treffer	rel. Häuf.	Texte	von	bis	Quelle
19	9.8 pMW	18	2002	2006	Reden und Interviews
26	6.5 pMW	16	2006	2007	VDI Nachrichten
28	8.2 pMW	19	2008	2009	VDI nachrichten
65	11.5 pMW	51	1998	2009	spektrumdirekt
292	15.8 pMW	256	2007	2009	Burgenländische Volkszeitung
340	37.8 pMW	287	1999	1999	Oberösterreichische Nachrichten
588	12.5 pMW	453	2006	2009	Nürnberger Nachrichten
609	8.9 pMW	465	2006	2009	Die Südostschweiz
624	17.8 pMW	429	2007	2009	Hannoversche Allgemeine
669	23.2 pMW	506	2006	2009	Hamburger Morgenpost
923	21.4 pMW	632	2006	2009	Nürnberger Zeitung
1.029	19.0 pMW	691	2005	2005	Wikipedia
1.097	6.4 pMW	893	1997	2009	St. Galler Tagblatt
1.196	10.2 pMW	941	2005	2009	Braunschweiger Zeitung
1.353	15.9 pMW	999	2006	2009	Mannheimer Morgen

Oben links können Sie die Ansicht verändern. Sie können z.B. auch nach Jahr, Jahrzehnt oder Textsorten sortieren. Die Ansicht im Reiter „Ergebnisse“ ist interessant, da wir hier Gesamtfrequenzen ablesen können. In unserem Beispiel aber sagen diese Frequenzen wenig aus, da sich noch zu viele Fehltreffer (wie eben z.B. *Schwarzenegger*) unter den Ergebnissen befinden. Wir sind zunächst an den Belegen selbst interessiert, die wir über die **KWIC-Ansicht** im entsprechenden Reiter ansehen:

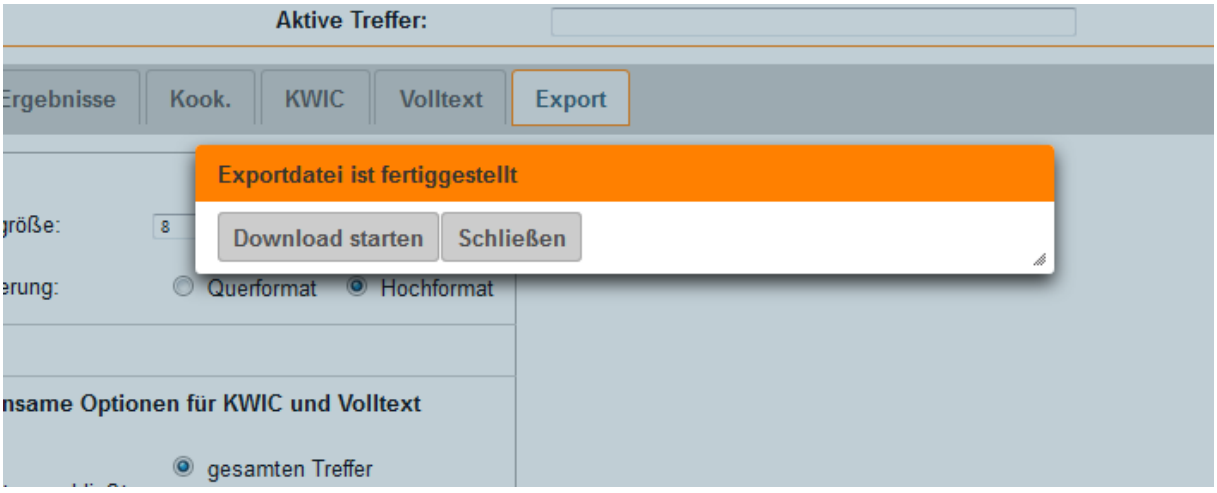


KWIC steht für **Key Word In Context** - und genau das sehen wir hier auch. Wieder können wir nach verschiedenen Kriterien sortieren, z.B. chronologisch, wie im obigen Screenshot. Beachten Sie, dass nur die Sortierungen, die Sie im KWIC-Reiter schon vorgenommen haben, auch für den Export verfügbar sind.
Genau diesem - dem Export also - wenden wir uns nun zu, indem wir wieder auf den entsprechenden Reiter wechseln.



Beachten Sie, dass maximal 10.000 Belege exportiert werden. Unsere Suchanfrage hat über 15.000 Ergebnisse erzielt. Für unser Beispiel leben wir damit, dass COSMAS nun eine Zufallsauswahl für uns trifft - methodisch noch sauberer wäre es allerdings, ein kleineres Korpus bzw. mehrere kleinere Korpora zu definieren.
Wir lassen uns die Belege als ASCII-Dokument ausgeben, das wir anschließend herunterladen und **mit Notepad** öffnen können.

7. Arbeit mit dem KWIC-Export



Wenn Sie die gleichen Einstellungen verwendet haben wie ich (vgl. Screenshot auf der vorigen Seite), erhalten Sie ein Textdokument, das aus drei Teilen besteht[1]:

1.) Header mit der Suchanfrage und Ergebnisübersicht

```
© Institut für Deutsche Sprache, Mannheim
COSMAS II-Server, C2API-Version 4.5.6 - 09. Oktober 2014
_____________________________________________________________

Datum         : Mittwoch, den 21. Januar 2015, 15:32:15
Korpus        : TAGGED-C-öffentlich - alle öffentlichen Korpora des Archivs TAGGED-C
Archiv-Release: Deutsches Referenzkorpus (DeReKo-2010-II)
Suchanfrage   : Abstand (+0w Max)
                 &#9492;ODER
                  &#9492;Schwarz*er
                  &#9492;Schwarz*erin
                  &#9492;Schwarz*erinnen
                 &#9492;ELEM(ANA='N')
Suchoptionen  : Ei+Ri+Di, Flex+Sonst
Ergebnis      : 15.194 Treffer


Ergebnisübersicht

Jahrzehntansicht, 2 Jahrzehnte, nach »Treffer« aufsteigend sortiert.

Anz Treffer (kumulativ) rel. Häuf. Anz. Texte : Jahrzehnt

      2.359       2.359 <rend 153,153,153> 28.26 pMW</rend>      1.756 : 1990-1999
     12.835      15.194 <rend 153,153,153> 13.66 pMW</rend>      9.706 : 2000-2009

     15.194      15.194  14.85 pMW     11.462 : 2 Jahrzehnte


KWIC (chronologisch sortiert)

Anz. Treffer           : 15.194
Anz. exportierte Zeilen: 10.000 (interne Begrenzung)
Angezeigter Kontext    : 20 Wörter links, 20 Wörter rechts
Kontext umschließt     : gesamten Treffer
_____________________________________________________________

```

2.) KWIC

```
1997
A97   Sicherheitsreferent und Landesrat Erich <B>Schwärzler</> am Mittwoch an einem Pressegespräch
A97   in Bregenz. Schwerverbrecher abschieben <B>Schwärzler</> bekräftigte die Auffassung, die
A97   erfunden hat, der nichts gemein hat mit <B>Schwarzenegger</> oder Stallone. Mich interessiert
A97   erfunden hat, der nichts gemein hat mit <B>Schwarzenegger</> oder Stallone. Mich interessiert
A97  ihnen anzutreffen, ausser denen, die wie <B>Schwarzfahrer</> entdeckt werden und nun auf eine
A97       kündigte Marent an. Landesrat Erich <B>Schwärzler</> will nicht nur, dass die
A97    Fürstentums Liechtenstein zu sehen. Im <B>Schwarzenberger</> Gemeindemuseum wird zudem (23.
A97      - für Monika Schumacher kommt Robert <B>Schwarzer.</> Amriswil: Mit Gegenständen aus der
A97  ist vielleicht ein Zeichen dafür: Robert <B>Schwarzer</> (45) will kein "Angepasster" sein.
A97 dasein Der vierfache Familienvater Robert <B>Schwarzer</> ist in Arbons Altstadt aufgewachsen.
A97   er sich damals, sondern auch politisch: <B>Schwarzer</> war bereits in den Jahren 1980 bis
A97        nachgerückt. 1980 allerdings hatte <B>Schwarzer</> einstweilen «genug von der Politik» -
A97 Wahlen zwei Sitze an die «Arboner Liste». <B>Schwarzer</> wurde ein "Opfer" des Proporzes,
A97  plus Stadtammann formiert werden. Robert <B>Schwarzer</> könnte sich durchaus vorstellen, sich
A97            mehr Heute "wehrt" sich Robert <B>Schwarzer</> hauptberuflich für die Anliegen der
A97   gemütliches Beisammensein im Restaurant <B>Schwarzer</> Adler. Der Vorstand und die
A97 im Rat Der neue Ortsverwaltungsrat Robert <B>Schwarzer,</> der ab 1. Juni Monika Schumacher
A97  gestern hiess es fälschlicherweise, dass <B>Schwarzer</> von 1980 bis 1993 OV-Rat gewesen
A97    gibt, schmeissen sie Cola-Flaschen mit <B>Schwarzpulver</> ins Wasser. Auch Fangmethoden mit
A97  der erste Präsident der Ortspartei, Reto <B>Schwarzenberger,</> am Gründungsfest. Diese sei
A97    Sommer ändern. Im Vorstand sind, neben <B>Schwarzenberger,</> Franziska Frohofer als Akuarin
1998
A98            trainieren regelmässig sieben <B>Schwarzgurtträger,</> die ihre Erfahrung und ihr
A98    dabei die zwei extremen Positionen der <B>Schwarzmaler</> und derjenigen, welche die Welt
A98    in Balgach sind lebensmüde unsichtbare <B>"Schwarzfahrer"</> unterwegs. In Heerbrugg steht
A98            trainieren regelmässig sieben <B>Schwarzgurtträger,</> die ihre Erfahrung und ihr
A98      Bass und Schlagzeug war dagegen eher <B>Schwarzenegger-</>Performance als Jazz-Konzert.
A98  findet in der Bürgerstube des Restaurant <B>Schwarzer</> Adler die Hauptversammlung des
A98    die Hauptversammlung des Kulturkreises <B>Schwarzer</> Adler statt. Es wird auch das neue
A98   23. Januar, um 19.30 Uhr ins Restaurant <B>Schwarzer</> Adler, Gossau ein. Nebst den
A98    ermögliche die Benützung aller Loipen. <B>Schwarzfahrer</> belasten den Verein, der eine
A98 Jagd auf Sozialbetrüger, Asylbewerber und <B>Schwarzarbeiter</> noch der Verweis auf
A98    die Hauptversammlung des Kulturkreises <B>Schwarzer</> Adler statt. Dabei werden die
A98 findet in der Bürgerstube des Restaurants <B>Schwarzer</> Adler in Gossau die Hauptversammlung
A98            Eindruck hinterliess auch die <B>Schwarzenbacher</> Guggenmusik "Thurvögel". Ihre
```

[1] Welche dieser Teile tatsächlich vorhanden sind, kann allerdings variieren, je nachdem, welche Häkchen Sie im Exportfenster gesetzt haben. Soweit ich weiß, werden aber defaultmäßig die drei hier genannten exportiert.

3.) Volltext

```
10044  NUN08        mit Bezirksvorsitzenden Wittibald <B>Schwarzer</> zu ihrem 60-jährigen Jubiläum in den
10045  NUZ08   von Felix Mitterer und wird von Xaver <B>Schwarzenberger</> in Szene gesetzt. Die Bären im
10046  RHZ08   Tanzproduktion, aber auch Sprech- und <B>Schwarzlichttheater.</> Die Schauspieler stammen
10047  RHZ08   der Kreisverwaltung in Simmern, Armin <B>Schwarzfischer,</> Tel. 06761/828 53, oder Simon
10048  RHZ08         zu besprechen", erläutert Armin <B>Schwarzfischer.</> Gefördert werden private
10049  RHZ08  eine Online-Leerstandsbörse auf. Armin <B>Schwarzfischer</> vermutet, dass nicht einmal so
10050
10051
10052  Belege (chronologisch sortiert)
10053
10054  Anz. Treffer           : 15.194
10055  Anz. exportierte Belege: 10.000 (interne Begrenzung)
10056  Angezeigter Kontext    : 0 Absätze links, 0 Absätze rechts
10057  Kontext umschließt     : gesamten Treffer
10058  ____________________________________________________________________
10059
10060  Unabhängig<A @NH> davon<ADV @ADVL> sei<V SUB PRES @MAIN> aus<PREP @PREMARK> sicherheitspolitischer<A @PREMOD> Sicht<N @NH> die<DET @PREMOD> Möglichkeit<N @NH>
       grenzüberschreitender<A @PREMOD> Fahndung<N @NH> «ein<DET @PREMOD> Gebot<N @NH> der<DET @PREMOD> Stunde<N @NH>», betonte<V IND PAST @MAIN> Sicherheitsreferent<N @NH>
       und<CC @CC> Landesrat<N @NH> Erich<N Prop @PREMOD> <B>Schwärzler</><N Prop @NH> am<PREP @PREMARK> Mittwoch<N @NH> an<PREP @PREMARK> einem<DET @PREMOD>
       Pressegespräch<N @NH> mit<PREP @POSTMOD> Sicherheitsdirektor<N @NH> Elmar<N Prop @PREMOD> Marent,<N Prop @NH> Landespolizeikommandant<N @PREMOD> Brigadier<N @NH>
       Manfred<N Prop @PREMOD> Bliem<N Prop @NH> und<CC @CC> Zollwache-Vorstand<N @NH> Oberst<N @NH> Josef<N Prop @PREMOD> Ebner<N Prop @NH> in<PREP @PREMARK> Bregenz.<N
       Prop @NH> (A97/APR.00296 St. Galler Tagblatt, 24.04.1997, Ressort: TB-THG (Abk.); Verhandlungen mit der Schweiz)
10061
10062  <B>Schwärzler</><N Prop @NH> bekräftigte<V IND PAST @MAIN> die<DET @PREMOD> Auffassung,<N @NH> die<DET @PREMOD> positive<A @PREMOD> Integration<N @NH> von<PREP
       @POSTMOD> Ausländern<N PL @NH> komme<V SUB PRES @MAIN> für<PREP @PREMARK> Schwerverbrecher<N @NH> nicht<ADV @ADVL> in<PREP @PREMARK> Frage.<N @NH> Hier<ADV @ADVL>
       müsse<V SUB PRES @MAIN> entgegen<PREP @PREMARK> den<DET @PREMOD> geplanten<A @PREMOD> Gesetzen<N PL @NH> die<DET @PREMOD> unbürokratische<A @PREMOD> Abschiebung<N
       @NH> von<PREP @POSTMOD> Gewalttätern<N PL @NH> möglich<A @NH> sein.<V INF @MAIN> (A97/APR.00296 St. Galler Tagblatt, 24.04.1997, Ressort: TB-THG (Abk.); Verhandlungen
       mit der Schweiz)
10063
10064  Leon:<N Prop @NH> Es<PRON @NH> mag<V IND PRES @AUX> daran<ADV @ADVL> liegen,<V INF @MAIN> dass<CS @PREMARK> eine<DET @PREMOD> Frau<N @NH> diesen<PRON @PREMOD>
       Helden<N @NH> erfunden<V PCP PERF @MAIN> hat,<V IND PRES @MAIN> der<PRON @NH> nichts<PRON @PREMOD> gemein<A @NH> hat<V IND PRES @MAIN> mit<ADV @ADVL>
       <B>Schwarzenegger</><N Prop @NH> oder<CC @CC> Stallone.<N Prop @NH> Mich<PRON @NH> interessiert<V IND PRES @MAIN> das<PRON @NH> Machotum<N @NH> nicht.<ADV @ADVL>
       Ein<DET @PREMOD> Mann<N @NH> kann<V IND PRES @MAIN> zivilisiert,<V PCP PERF @MAIN> gebildet,<V PCP PERF @MAIN> feinfühlig<A @NH> und<CC @CC> dennoch<ADV @ADVL>
       sehr<ADV @PREMOD> männlich<A @NH> sein<V INF @MAIN> wie<PREP @PREMARK> Brunetti,<N Prop @NH> das<PRON @NH> ist<V IND PRES @MAIN> es<PRON @NH> vermutlich,<A @NH>
       was<PRON @NH> den<DET @PREMOD> Leserinnen<N PL @NH> gefällt.<V IND PRES @MAIN> Er<PRON @NH> ähnelt<V IND PRES @MAIN> übrigens<ADV @ADVL> Inspektor<N @NH> Wexford,<N
       Prop @NH> dem<DET @PREMOD> kultivierten<A @PREMOD> Helden<N @NH> der<DET @PREMOD> Krimis<N PL @NH> der<DET @PREMOD> Engländerin<N @PREMOD> Ruth<N Prop @PREMOD>
```

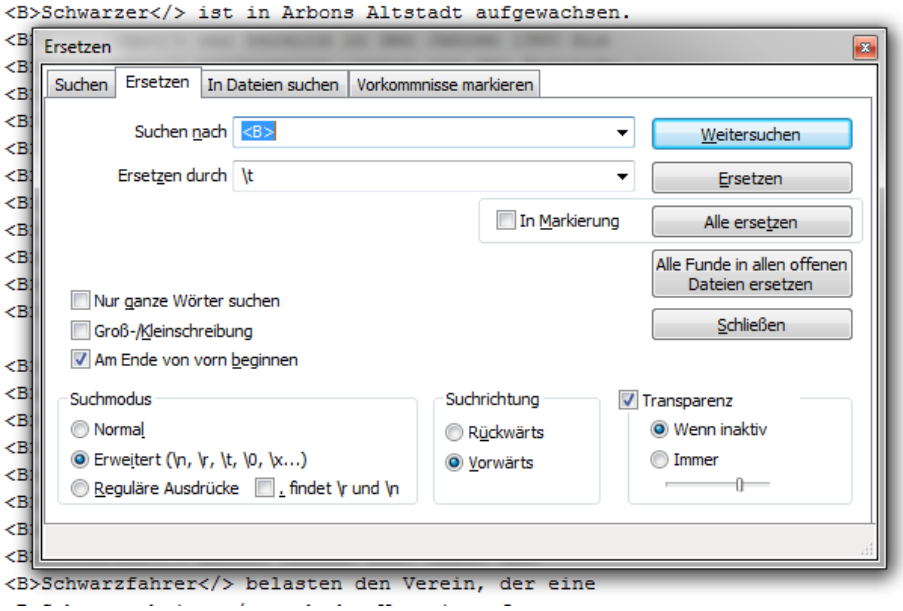
Der Volltext sieht hier etwas kryptisch aus, da die Annotationen mit eingeblendet sind - das lässt sich in den Exportoptionen auch einfach ändern. Uns interessiert aber zunächst nur der KWIC-Teil.

7.1 KWIC-Export in ein Spreadsheet überführen: Der manuelle Weg

Der KWIC-Export lässt sich mit dem R-Paket „concordances“ relativ einfach und bequem in R einlesen und dann wieder als CSV-Spreadsheet exportieren (s. Tutorial zu „concordances“). In diesem Abschnitt zeige ich dennoch auf, wie man mit Hilfe von Notepad++ das gleiche Ergebnis manuell erzielen kann. Wenn Sie lieber den einfachen Weg gehen möchten, können Sie einfach zu 7.2 springen.
Zunächst copy&pasten wir den KWIC-Teil der Exportdatei (und nur diesen) in ein neues Notepad-Dokument.

```
A97   Sicherheitsreferent und Landesrat Erich <B>Schwärzler</> am Mittwoch an einem Pressegespräch
A97   in Bregenz. Schwerverbrecher abschieben <B>Schwärzler</> bekräftigte die Auffassung, die
A97   erfunden hat, der nichts gemein hat mit <B>Schwarzenegger</> oder Stallone. Mich interessiert
A97   erfunden hat, der nichts gemein hat mit <B>Schwarzenegger</> oder Stallone. Mich interessiert
A97  ihnen anzutreffen, ausser denen, die wie <B>Schwarzfahrer</> entdeckt werden und nun auf eine
A97        kündigte Marent an. Landesrat Erich <B>Schwärzler</> will nicht nur, dass die
A97     Fürstentums Liechtenstein zu sehen. Im <B>Schwarzenberger</> Gemeindemuseum wird zudem (23.
A97       - für Monika Schumacher kommt Robert <B>Schwarzer.</> Amriswil: Mit Gegenständen aus der
A97  ist vielleicht ein Zeichen dafür: Robert <B>Schwarzer</> (45) will kein "Angepasster" sein.
A97 dasein Der vierfache Familienvater Robert <B>Schwarzer</> ist in Arbons Altstadt aufgewachsen.
A97   er sich damals, sondern auch politisch: <B>Schwarzer</> war bereits in den Jahren 1980 bis
A97        nachgerückt. 1980 allerdings hatte <B>Schwarzer</> einstweilen «genug von der Politik» -
A97 Wahlen zwei Sitze an die «Arboner Liste». <B>Schwarzer</> wurde ein "Opfer" des Proporzes,
A97  plus Stadtammann formiert werden. Robert <B>Schwarzer</> könnte sich durchaus vorstellen, sich
A97            mehr Heute "wehrt" sich Robert <B>Schwarzer</> hauptberuflich für die Anliegen der
A97   gemütliches Beisammensein im Restaurant <B>Schwarzer</> Adler. Der Vorstand und die
A97 im Rat Der neue Ortsverwaltungsrat Robert <B>Schwarzer,</> der ab 1. Juni Monika Schumacher
A97  gestern hiess es fälschlicherweise, dass <B>Schwarzer</> von 1980 bis 1993 OV-Rat gewesen
A97    gibt, schmeissen sie Cola-Flaschen mit <B>Schwarzpulver</> ins Wasser. Auch Fangmethoden mit
A97  der erste Präsident der Ortspartei, Reto <B>Schwarzenberger,</> am Gründungsfest. Diese sei
A97    Sommer ändern. Im Vorstand sind, neben <B>Schwarzenberger,</> Franziska Frohofer als Aktuarin
1998
A98            trainieren regelmässig sieben <B>Schwarzgurtträger,</> die ihre Erfahrung und ihr
A98    dabei die zwei extremen Positionen der <B>Schwarzmaler</> und derjenigen, welche die Welt
A98    in Balgach sind lebensmüde unsichtbare <B>"Schwarzfahrer"</> unterwegs. In Heerbrugg steht
A98            trainieren regelmässig sieben <B>Schwarzgurtträger,</> die ihre Erfahrung und ihr
A98      Bass und Schlagzeug war dagegen eher <B>Schwarzenegger-</>Performance als Jazz-Konzert.
A98  findet in der Bürgerstube des Restaurant <B>Schwarzer</> Adler die Hauptversammlung des
A98    die Hauptversammlung des Kulturkreises <B>Schwarzer</> Adler statt. Es wird auch das neue
A98   23. Januar, um 19.30 Uhr ins Restaurant <B>Schwarzer</> Adler, Gossau ein. Nebst den
A98    ermögliche die Benützung aller Loipen. <B>Schwarzfahrer</> belasten den Verein, der eine
A98 Jagd auf Sozialbetrüger, Asylbewerber und <B>Schwarzarbeiter</> noch der Verweis auf
A98    die Hauptversammlung des Kulturkreises <B>Schwarzer</> Adler statt. Dabei werden die
A98 findet in der Bürgerstube des Restaurants <B>Schwarzer</> Adler in Gossau die Hauptversammlung
A98            Eindruck hinterliess auch die <B>Schwarzenbacher</> Guggenmusik "Thurvögel". Ihre
A98          Der zuständige Landesrat Erich <B>Schwärzler</> sei im Kontakt mit dem
```

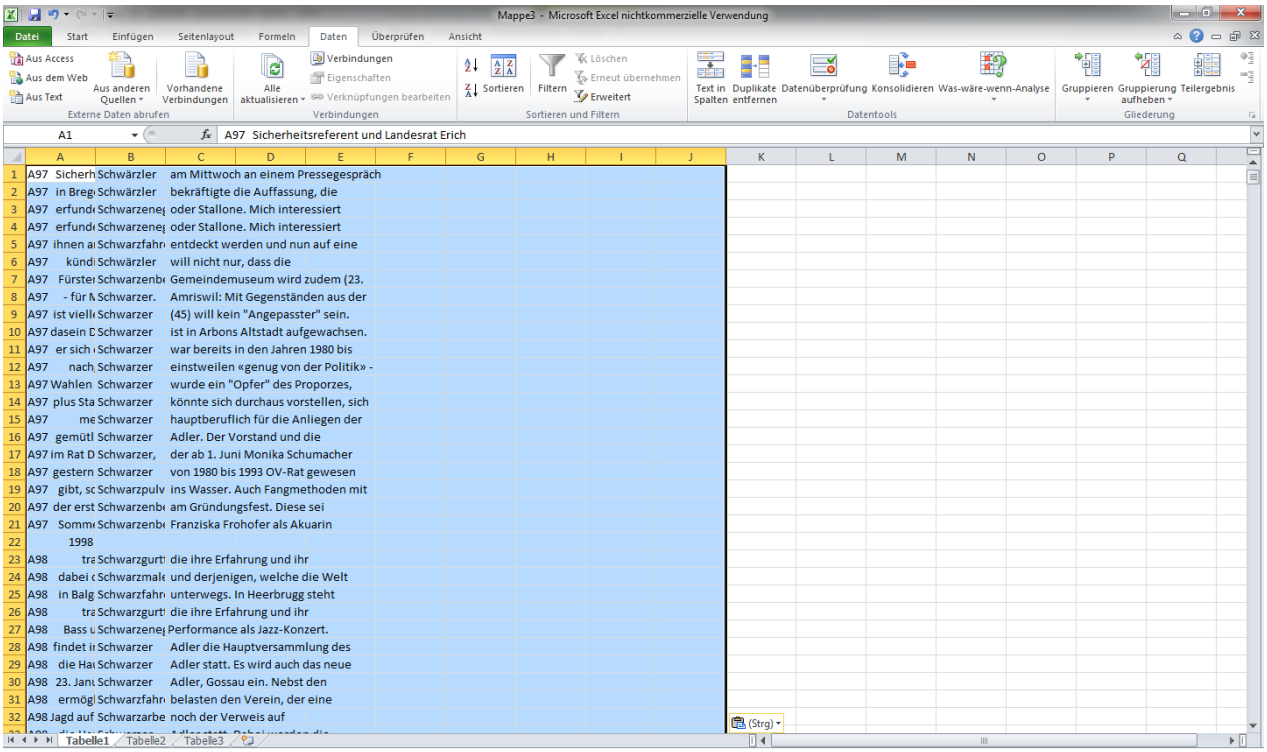
Wie Sie sehen, ist das **Keyword** in <B> und </> eingeschlossen. Um die Daten in ein Tabellenkalkulationsprogramm (z.B. Excel) zu exportieren, **ersetzen** wir diese beiden Marker durch **Tabs**. Über *Suchen>Ersetzen* oder einfach Strg+H gelangen wir in das Ersetzungsfenster.



Stellen Sie sicher, dass als Suchmodus „erweitert“ ausgewählt ist: Nur so können wir \t als Code für Tabs verwenden. Wir ersetzen nun alle <B> durch \t und wiederholen dies für </> (jeweils mit „Alle ersetzen“, dritter Button von oben). Die Datei sieht nun so aus:

```
schwarz_cosmas.TXT   new 1
1  A97   Sicherheitsreferent und Landesrat Erich   Schwärzler   am Mittwoch an einem Pressegespräch
2  A97   in Bregenz. Schwerverbrecher abschieben   Schwärzler   bekräftigte die Auffassung, die
3  A97   erfunden hat, der nichts gemein hat mit   Schwarzenegger   oder Stallone. Mich interessiert
4  A97   erfunden hat, der nichts gemein hat mit   Schwarzenegger   oder Stallone. Mich interessiert
5  A97  ihnen anzutreffen, ausser denen, die wie   Schwarzfahrer    entdeckt werden und nun auf eine
6  A97       kündigte Marent an. Landesrat Erich   Schwärzler   will nicht nur, dass die
7  A97    Fürstentums Liechtenstein zu sehen. Im   Schwarzenberger  Gemeindemuseum wird zudem (23.
8  A97      - für Monika Schumacher kommt Robert   Schwarzer.   Amriswil: Mit Gegenständen aus der
9  A97  ist vielleicht ein Zeichen dafür: Robert   Schwarzer    (45) will kein "Angepasster" sein.
10 A97 dasein Der vierfache Familienvater Robert   Schwarzer    ist in Arbons Altstadt aufgewachsen.
11 A97   er sich damals, sondern auch politisch:   Schwarzer    war bereits in den Jahren 1980 bis
12 A97        nachgerückt. 1980 allerdings hatte   Schwarzer    einstweilen «genug von der Politik» -
13 A97 Wahlen zwei Sitze an die «Arboner Liste».   Schwarzer    wurde ein "Opfer" des Proporzes,
14 A97  plus Stadtammann formiert werden. Robert   Schwarzer    könnte sich durchaus vorstellen, sich
15 A97            mehr Heute "wehrt" sich Robert   Schwarzer    hauptberuflich für die Anliegen der
16 A97   gemütliches Beisammensein im Restaurant   Schwarzer    Adler. Der Vorstand und die
17 A97 im Rat Der neue Ortsverwaltungsrat Robert   Schwarzer,   der ab 1. Juni Monika Schumacher
18 A97  gestern hiess es fälschlicherweise, dass   Schwarzer    von 1980 bis 1993 OV-Rat gewesen
19 A97    gibt, schmeissen sie Cola-Flaschen mit   Schwarzpulver    ins Wasser. Auch Fangmethoden mit
20 A97  der erste Präsident der Ortspartei, Reto   Schwarzenberger,     am Gründungsfest. Diese sei
21 A97    Sommer ändern. Im Vorstand sind, neben   Schwarzenberger,     Franziska Frohofer als Akuarin
22 1998
23 A98             trainieren regelmässig sieben   Schwarzgurtträger,   die ihre Erfahrung und ihr
24 A98    dabei die zwei extremen Positionen der   Schwarzmaler     und derjenigen, welche die Welt
25 A98    in Balgach sind lebensmüde unsichtbare   "Schwarzfahrer"  unterwegs. In Heerbrugg steht
26 A98             trainieren regelmässig sieben   Schwarzgurtträger,   die ihre Erfahrung und ihr
27 A98      Bass und Schlagzeug war dagegen eher   Schwarzenegger- Performance als Jazz-Konzert.
28 A98  findet in der Bürgerstube des Restaurant   Schwarzer    Adler die Hauptversammlung des
29 A98    die Hauptversammlung des Kulturkreises   Schwarzer    Adler statt. Es wird auch das neue
30 A98   23. Januar, um 19.30 Uhr ins Restaurant   Schwarzer    Adler, Gossau ein. Nebst den
31 A98    ermögliche die Benützung aller Loipen.   Schwarzfahrer    belasten den Verein, der eine
```

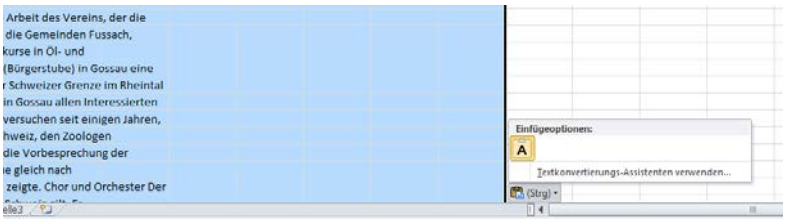
Diesen Text können wir direkt in Excel copy&pasten. Zunächst interpretiert Excel die Daten jedoch falsch, wie Sie am folgenden Screenshot sehen. Versuchen Sie, den Fehler zunächst selbst zu erkennen, bevor Sie weiterlesen.



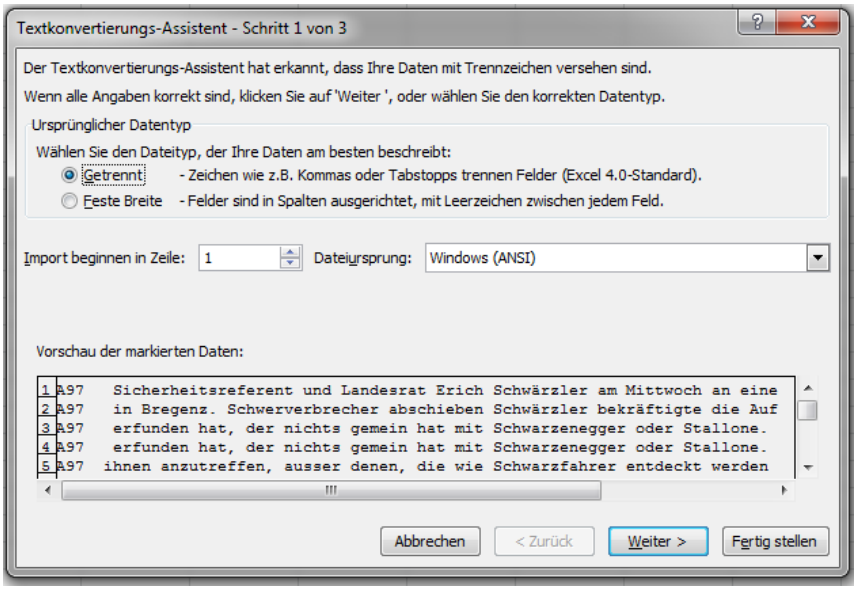
Vielleicht haben Sie es erkannt: Eigentlich sollten die Daten genau **drei** Spalten füllen. Hier verteilen sich die Daten aber auf bis zu **zehn** Spalten (im dargestellten Ausschnitt sind zwar die letzten sieben allesamt leer, aber es gibt durchaus Zeilen, bei denen sich in den Spalten D bis J Text findet.) Woran liegt das?

Eine mögliche Erklärung wäre, dass in den Rohdaten bereits Tabs vorhanden waren. Das ist bei COSMAS meines Wissens nie der Fall, aber grundsätzlich ist es eine Möglichkeit, die man für zukünftige Arbeiten auch mit anderen Korpora im Hinterkopf behalten sollte. Hier ist das Problem jedoch ein anderes: Die COSMAS-Texte enthalten **Anführungszeichen.** Diese behandelt Excel als **Textqualifizierer**, also als Marker, die Excel mitteilen, dass es die ent-

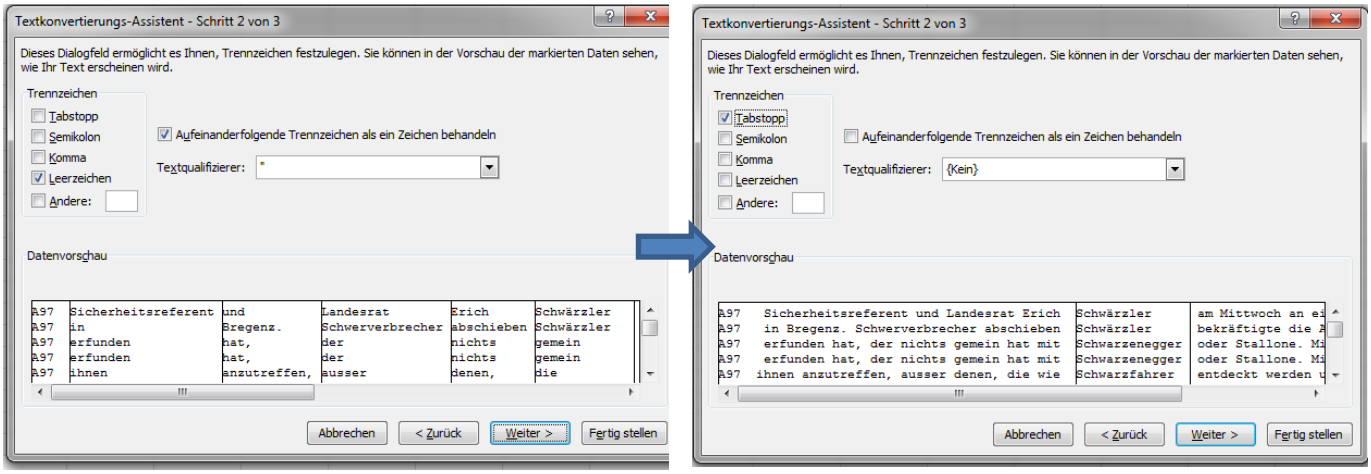
sprechenden Daten als Text behandeln soll. So kommt es dazu, dass teilweise die Daten zwischen zwei Anführungszeichen als ein zusammenhängender Text behandelt werden, was zu Chaos in den Daten führt. Um dies zu verhindern, klicken wir gleich nach dem Einfügen auf das kleine Kästchen neben der Markierung, in dem „{Strg}“ steht (dieses verschwindet wieder, sobald wir die Markierung aufheben, deshalb gleich draufklicken!).



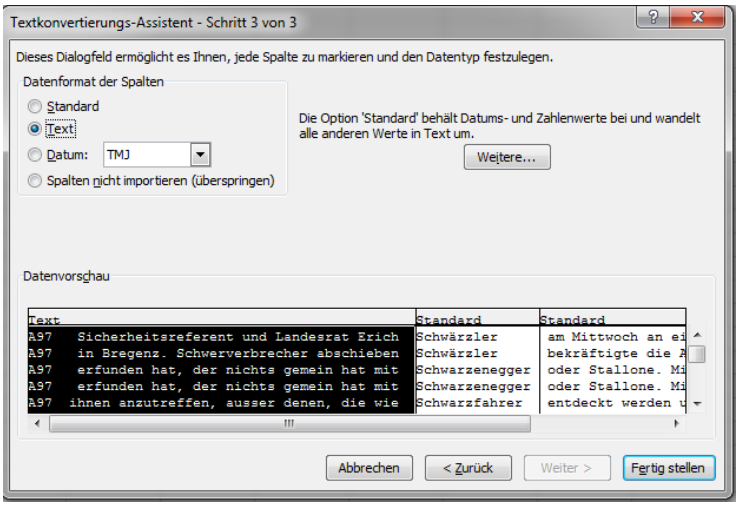
Hier wählen wir die Option „Textkonvertierungs-Assistenten verwenden“. Nun können wir Excel in drei Schritten mitteilen, wie es die Daten interpretieren soll:



Im ersten Schritt ist die von uns gewollte Option „Getrennt“ bereits angewählt, wir können also einfach auf „Weiter“ klicken.



Im zweiten Schritt wählen wir **Tabstopp** als Trennzeichen und - ganz wichtig - deaktivieren die Erkennung von Anführungszeichen als Textqualifizierer, indem wir als **Textqualifizierer** *{kein}* auswählen.



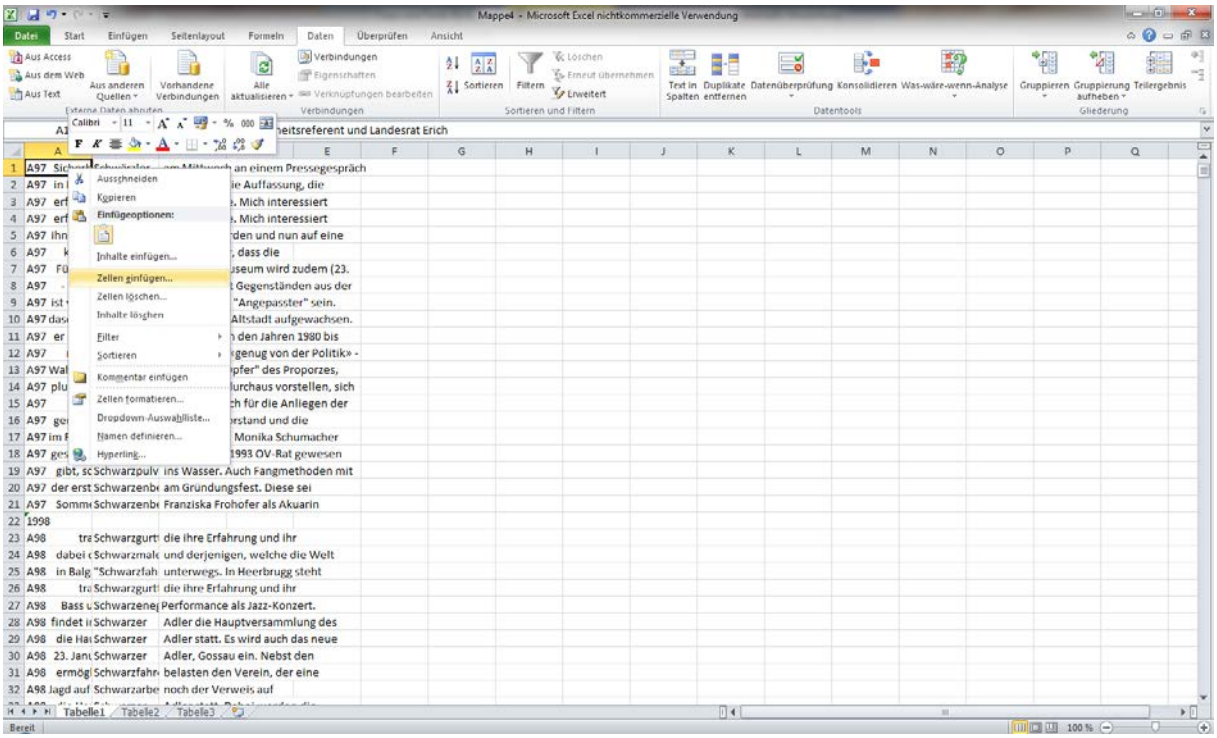
Im dritten Schritt schließlich können wir Excel noch mitteilen, dass es sich bei unseren Daten um **Text** handelt (nicht etwa um Zahlen oder um Kalenderdaten). Abschließend klicken wir auf **Fertig stellen.**

7.2 KWIC-Export in ein Spreadsheet mit „concordances“

Mit dem R-Paket „concordances“ lässt sich der COSMAS-Export unschwer in ein CSV-Spreadsheet überführen. Folgen Sie dazu einfach der Anleitung im Tutorial zu „concordances“. Um das so entstandene CSV-Sheet zu öffnen, folgen Sie der Anleitung im Tutorial „01-Grundlegendes“.

8. Arbeit mit dem Excel-Dokument

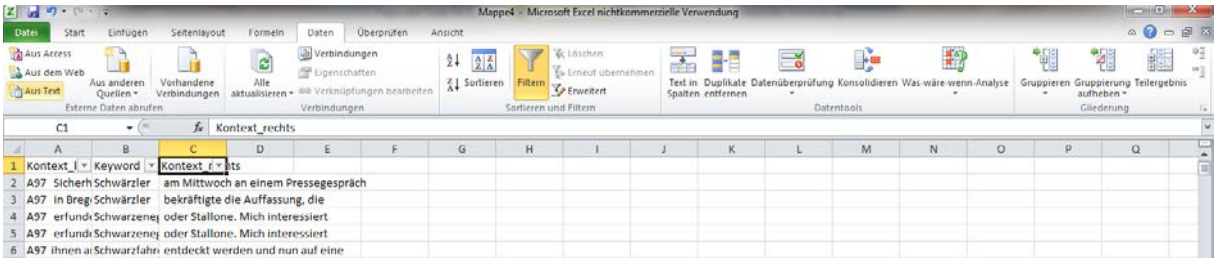
Geschafft: Wir haben unsere **Konkordanz**, also unsere Belegsammlung, erfolgreich in Excel eingefügt. Im Tutorial „06-Tabellen und Grafiken mit Excel und Calc“ gehen wir noch ausführlich auf die Frage ein, wie man Daten in Tabellenkalkulationsprogammen tabellarisch darstellen oder auch in Diagrammen visualisieren kann; hier seien schon einige Schritte vorweggenommen. Das meiste lässt sich übrigens 1:1 auf LibreOffice Calc übertragen.
Um die Daten im Excel-Dokument nach Keyword sortieren zu können, fügen wir zunächst eine Überschriftenzeile ein: Rechtsklick in der ersten Zeile > Zellen einfügen > **Ganze Zeile.**



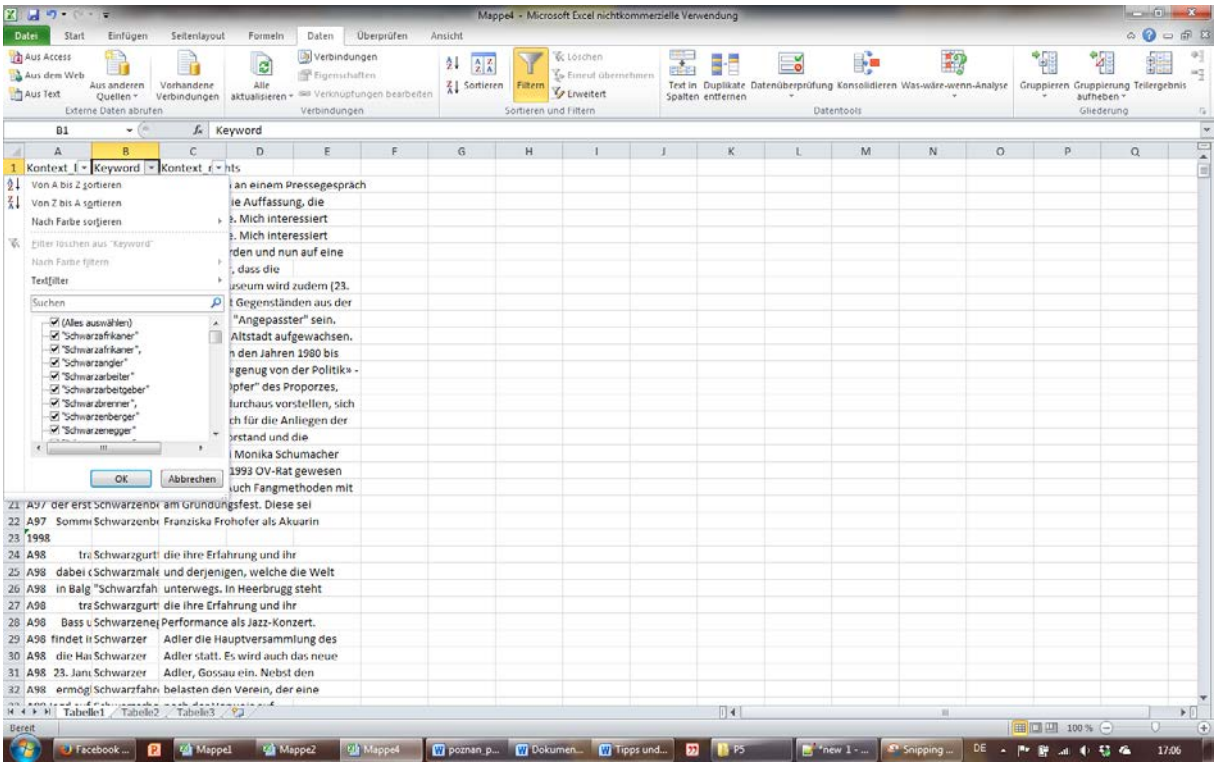
In die neu entstandene Zeile tragen wir nun die Überschriften für die einzelnen Spalten ein, z.B. Kontext_links, Keyword, Kontext_rechts.

	A	B	C	D	E
1	Kontext_link	Keyword	Kontext_rechts		
2	A97 Sicherh	Schwärzler	am Mittwoch an einem Pressegespräch		
3	A97 in Breg	Schwärzler	bekräftigte die Auffassung, die		
4	A97 erfund	Schwarzeneg	oder Stallone. Mich interessiert		
5	A97 erfund	Schwarzeneg	oder Stallone. Mich interessiert		
6	A97 ihnen a	Schwarzfahr	entdeckt werden und nun auf eine		
7	A97 kündi	Schwärzler	will nicht nur, dass die		
8	A97 Fürster	Schwarzenbe	Gemeindemuseum wird zudem (23.		
9	A97 - für N	Schwarzer.	Amriswil: Mit Gegenständen aus der		
10	A97 ist vielle	Schwarzer	(45) will kein "Angepasster" sein.		
11	A97 dasein D	Schwarzer	ist in Arbons Altstadt aufgewachsen.		
12	A97 er sich	Schwarzer	war bereits in den Jahren 1980 bis		
13	A97 nach	Schwarzer	einstweilen «genug von der Politik» -		
14	A97 Wahlen	Schwarzer	wurde ein "Opfer" des Proporzes,		
15	A97 plus Sta	Schwarzer	könnte sich durchaus vorstellen, sich		
16	A97 me	Schwarzer	hauptberuflich für die Anliegen der		

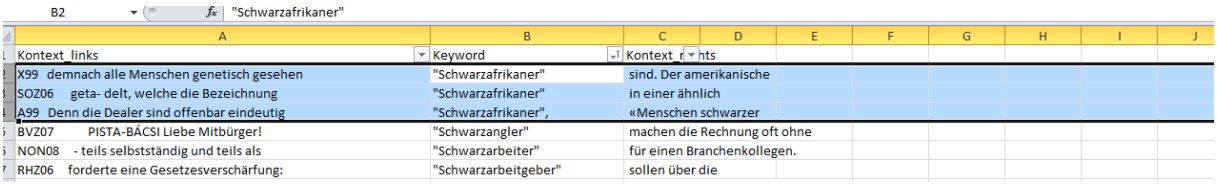
Wir lassen die erste Zeile markiert, gehen ganz oben auf den Reiter DATEN und dort auf **Filtern**.



Nun sehen Sie kleine Pfeilchen neben den Überschriften. Mit einem Klick auf das Pfeilchen neben **Keyword** können wir die Schlüsselwörter „Von A bis Z sortieren“:



Damit haben wir unsere Liste de facto nach **Types** sortiert und können Fehltreffer bequem aussortieren. Tipp: Mit **Strg + Leertaste** markieren Sie **die ganze Zeile**, mit Strg + - **löschen** Sie den markierten Bereich. Die ersten drei Belege in der hier vorliegenden Konkordanz lauten *Schwarzafrikaner* (das ist in Ihrer Konkordanz wegen anderer Zufallsauswahl evtl. anders). Das ist zwar eine Personenbezeichnung, aber trotz ihres rassistischen Gehalts und der damit einhergehenden negativen Konnotation keine produktive Verwendung von *Schwarz-*, die uns interessieren würde. Deshalb markieren wir die ersten drei Zeilen (mit Umschalt + Pfeiltasten können Sie mehrere Zeilen markieren) und dehnen die Markierung mit Hilfe der oben genannten Tastenkombination **Strg + Leertaste** dann auf die **gesamten Spalten** aus.

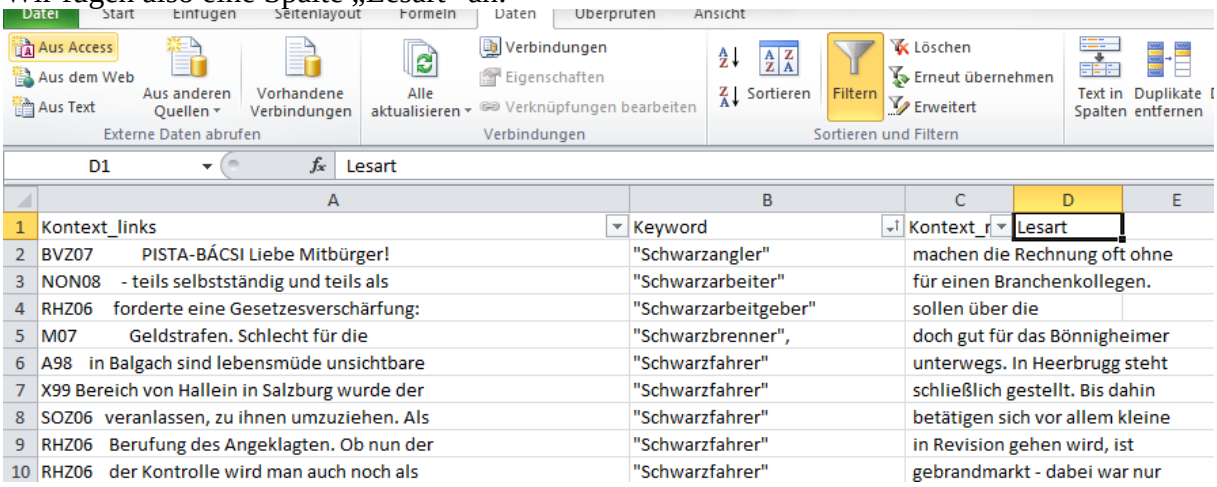


B2 "Schwarzafrikaner"

A	B	C	D	E	F	G	H	I	J
Kontext_links	Keyword	Kontext_r	hts						
X99 demnach alle Menschen genetisch gesehen	"Schwarzafrikaner"	sind. Der amerikanische							
SOZ06 geta- delt, welche die Bezeichnung	"Schwarzafrikaner"	in einer ähnlich							
A99 Denn die Dealer sind offenbar eindeutig	"Schwarzafrikaner",	«Menschen schwarzer							
BVZ07 PISTA-BÁCSI Liebe Mitbürger!	"Schwarzangler"	machen die Rechnung oft ohne							
NON08 - teils selbstständig und teils als	"Schwarzarbeiter"	für einen Branchenkollegen.							
RHZ06 forderte eine Gesetzesverschärfung:	"Schwarzarbeitgeber"	sollen über die							

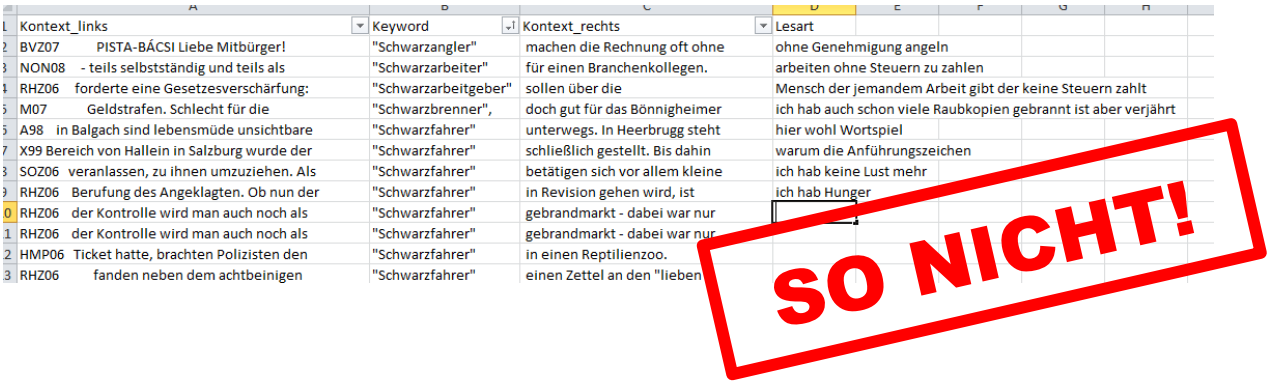
Nun können wir uns mit Hilfe von **Strg + -** der Fehltreffer entledigen. Die nächsten Belege, *Schwarzangler*, *Schwarzarbeiter* und *Schwarzarbeitgeber* sind hingegen genau die Verwendungen, die wir suchen. Bald kommen jedoch mit mehreren Belegen für *Schwarzer* wieder Fehltreffer. Hier können wir das eben skizzierte Vorgehen wiederholen, ebenso für alle anderen Fehltreffer, derer es hier sehr viele gibt. Umso besser, dass wir sie mit nur drei Tastenkombinationen schnell und effizient beseitigen können! Was übrigbleibt, ist eine recht überschaubare Belegsammlung, die wir ggf. noch näher annotieren wollen. Zum Beispiel könnten wir in einer Spalte festhalten, wie das *Schwarz-* jeweils zu interpretieren ist: Bezieht es sich auf eine „Kostenumgehung“ wie z.B. *Schwarzfahrer* oder *Schwarzpinkler* oder auf eine pes-

simistische Einstellung wie *Schwarzmaler* oder *Schwarzseher* (Letzteres jedoch ambig durch wortspielhaften Gebrauch z.B. in GEZ-Werbespots)?
Wir fügen also eine Spalte „Lesart“ an:



	A	B	C	D	E
1	Kontext_links	Keyword	Kontext_r	Lesart	
2	BVZ07 PISTA-BÁCSI Liebe Mitbürger!	"Schwarzangler"	machen die Rechnung oft ohne		
3	NON08 - teils selbstständig und teils als	"Schwarzarbeiter"	für einen Branchenkollegen.		
4	RHZ06 forderte eine Gesetzesverschärfung:	"Schwarzarbeitgeber"	sollen über die		
5	M07 Geldstrafen. Schlecht für die	"Schwarzbrenner",	doch gut für das Bönnigheimer		
6	A98 in Balgach sind lebensmüde unsichtbare	"Schwarzfahrer"	unterwegs. In Heerbrugg steht		
7	X99 Bereich von Hallein in Salzburg wurde der	"Schwarzfahrer"	schließlich gestellt. Bis dahin		
8	SOZ06 veranlassen, zu ihnen umzuziehen. Als	"Schwarzfahrer"	betätigen sich vor allem kleine		
9	RHZ06 Berufung des Angeklagten. Ob nun der	"Schwarzfahrer"	in Revision gehen wird, ist		
10	RHZ06 der Kontrolle wird man auch noch als	"Schwarzfahrer"	gebrandmarkt - dabei war nur		

Diese ist zunächst nicht gefiltert, aber indem wir den Filter oben deaktivieren und dann wieder aktivieren, stellen wir sicher, dass die gesamte Zeile gefiltert wird und wir auch die Lesart, wenn wir die entsprechende Spalte annotiert haben, alphabetisch sortiert und eben auch gefiltert werden kann.
Wichtig bei der Annotation, wenn Sie quantitativ an die Daten herangehen wollen: Arbeiten Sie mit wenigen, klar definierten **Variablenausprägungen**. Was heißt das? Ganz einfach: In unserem Fall ordnen Sie jeden Beleg z.B. der Ausprägung ILLEGAL oder der Ausprägung PESSIMISTISCH zu – auch wenn Sie bei einigen Belegen vielleicht Bauchschmerzen haben, sie einer der beiden Kategorien zuzuordnen.



	Kontext_links	Keyword	Kontext_rechts	Lesart
2	BVZ07 PISTA-BÁCSI Liebe Mitbürger!	"Schwarzangler"	machen die Rechnung oft ohne	ohne Genehmigung angeln
3	NON08 - teils selbstständig und teils als	"Schwarzarbeiter"	für einen Branchenkollegen.	arbeiten ohne Steuern zu zahlen
4	RHZ06 forderte eine Gesetzesverschärfung:	"Schwarzarbeitgeber"	sollen über die	Mensch der jemandem Arbeit gibt der keine Steuern zahlt
5	M07 Geldstrafen. Schlecht für die	"Schwarzbrenner",	doch gut für das Bönnigheimer	ich hab auch schon viele Raubkopien gebrannt ist aber verjährt
6	A98 in Balgach sind lebensmüde unsichtbare	"Schwarzfahrer"	unterwegs. In Heerbrugg steht	hier wohl Wortspiel
7	X99 Bereich von Hallein in Salzburg wurde der	"Schwarzfahrer"	schließlich gestellt. Bis dahin	warum die Anführungszeichen
8	SOZ06 veranlassen, zu ihnen umzuziehen. Als	"Schwarzfahrer"	betätigen sich vor allem kleine	ich hab keine Lust mehr
9	RHZ06 Berufung des Angeklagten. Ob nun der	"Schwarzfahrer"	in Revision gehen wird, ist	ich hab Hunger
10	RHZ06 der Kontrolle wird man auch noch als	"Schwarzfahrer"	gebrandmarkt - dabei war nur	
11	RHZ06 der Kontrolle wird man auch noch als	"Schwarzfahrer"	gebrandmarkt - dabei war nur	
12	HMP06 Ticket hatte, brachten Polizisten den	"Schwarzfahrer"	in einen Reptilienzoo.	
13	RHZ06 fanden neben dem achtbeinigen	"Schwarzfahrer"	einen Zettel an den "lieben	

Am Ende haben Sie dann – hoffentlich – ein übersichtliches Dokument mit Ihren Belegen und Annotationen, das Sie quantitativ auswerten können. Dafür gibt es in Excel ebenfalls einige hilfreiche Funktionen (mit denen ich mich aber weniger gut auskenne, da ich eher mit dem Statistikprogramm R arbeite). Ich hoffe jedoch, dass Ihnen dieses Tutorial den Einstieg in die Korpuslinguistik mit COSMAS II etwas einfacher gemacht hat. Für Verbesserungsvorschläge zu diesem Dokument bin ich natürlich jederzeit dankbar!