

Auch bei der Selektion: Ausdruckskraft durch Feinheiten

Im Zusammenhang mit Digitalisierung und Massendatenoperationen ist die Selektion eine der wichtigsten Operationen. Die Selektion für strukturierte Dokumente bietet weit mehr Möglichkeiten als die Selektion in flachen Tabellen des Relationalen Datenmodells. Hier gibt es lediglich die 2 Möglichkeiten. `SPALTENNAME=Wert` oder `SPALTENNAME1=SPALTENNAME2`. das Gleichheitszeichen kann noch durch andere Relationssymbole ersetzt werden.

Wir wollen einige Möglichkeiten von `o++o` anhand der Tabelle `ottos.tab` demonstrieren.

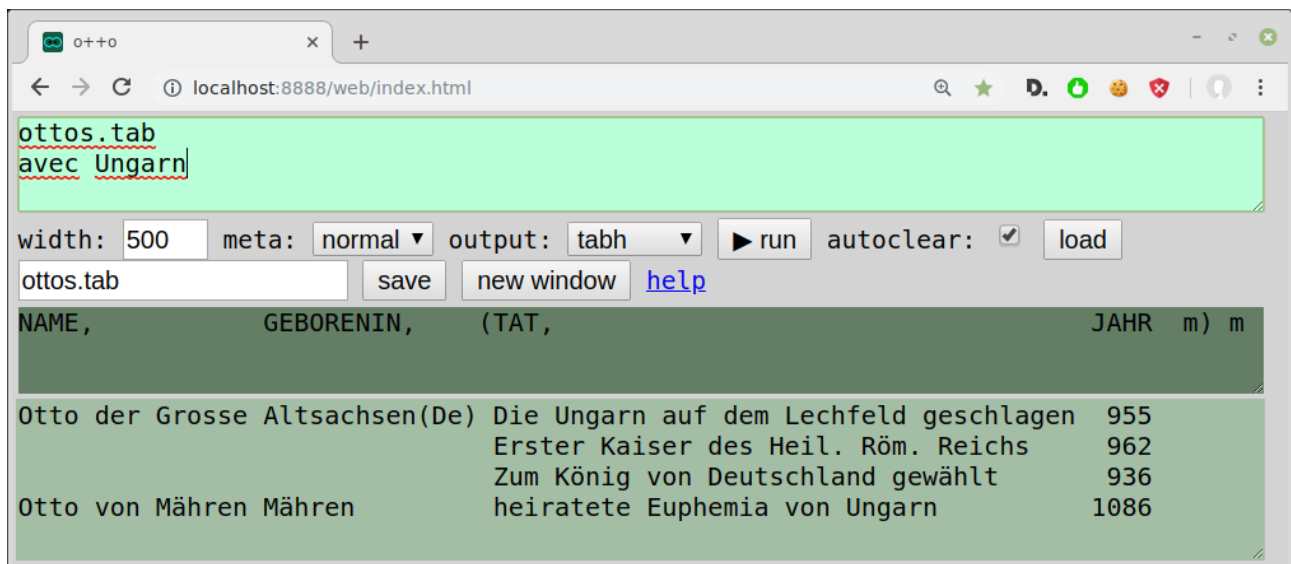
Die Selektion in `o++o` (`avec` bzw. `sans`) (mit bzw. ohne) hat zu nächst die vereinfachende Eigenschaft, dass sich der TT (TabmentTyp) und damit auch das Schema nicht ändert.

Selektion ohne Kollektionsspezifikation:

Gib alle Strupel, die das Wort Ungarn (irgendwo) enthalten.

`ottos.tab`

`avec Ungarn`



The screenshot shows the o++o web interface. The input field contains the query `ottos.tab` and `avec Ungarn`. The output is a table with the following data:

NAME,	GEBORENIN,	(TAT,	JAHR	m)	m
Otto der Grosse Altsachsen(De)	Die Ungarn auf dem Lechfeld geschlagen	Erster Kaiser des Heil. Röm. Reichs	955		
		Zum König von Deutschland gewählt	962		
			936		
Otto von Mähren Mähren	heiratete Euphemia von Ungarn		1086		

Die obige Selektion ist eine Abkürzung für:

`avec NAME! Ungarn in NAME tup`

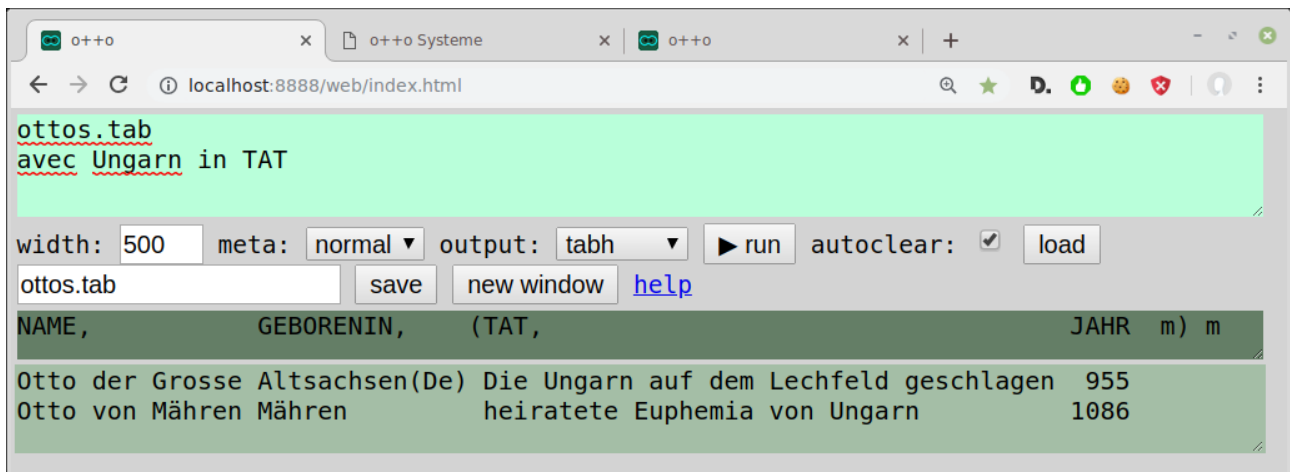
bzw.

`avec Ungarn in NAME tup`

und ist gleichwertig zu

`avec Ungarn in TATm`

Ungarn kommt nur in der Spalte TAT vor. Die Bedingung `Ungarn in TAT` unterscheidet sich trotzdem von den obigen, da jetzt in allen Kollektionen selektiert wird, die TAT enthalten. Das ist sowohl `(TAT, JAHR m)` als auch `(NAME, GEBORENIN, (TAT, JAHR m) m)`. `TATm` befindet sich eine Ebene höher als `TAT`, weswegen die Bedingung `Ungarn in TATm` lediglich in `(NAME, GEBORENIN, (TAT, JAHR m) m)` selektiert.



In der nächsten Anfrage besteht die Selektion lediglich aus 2 Wörtern:

#Gib alle Strupel, die die Worte Ungarn und König gleichzeitig enthalten.

ottos.tab
avec Ungarn König

Wenn ein avec-Teil aus mehr als einem (nicht Bool-) Wert besteht, kann seine Wirkung auch so definiert werden:

avec Wert1 Wert2 ... Wertn

ist das Gleiche wie die Hintereinanderanwendung mehrerer einfacherer Selektionen:

avec Wert1
avec Wert2
...
avec Wertn

Analog ist

sans Wert1 Wert2

gleichwertig zu

sans Wert1

sans Wert2

ottos.tab
sans Ungarn König

width: 500 meta: normal output: tabh run autoclear: load ottos.tab

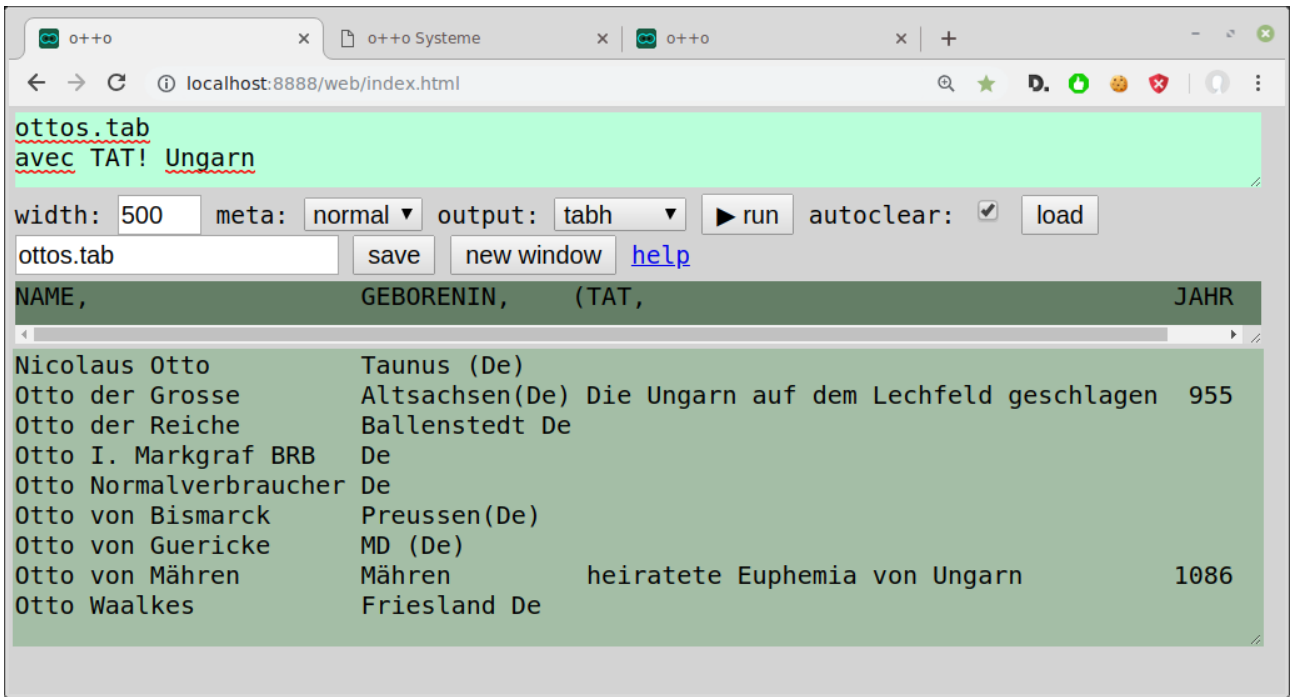
save new window help

NAME,	GEBORENIN,	(TAT,	JAHR m) m
Nicolaus Otto	Taunus (De)	Miterfinder des Ottomotors	1876
Otto der Reiche	Ballenstedt De	heiratete Eilika Billung von Sachsen	1094
		Herzog von Sachsen	1112
		Slavenkreuzzug	1108
		VaterAlbrechtsdesBären; GründerderMarkBrandenburg	1100
Otto I. Markgraf BRB	De	alleiniger Markgraf von Brandenburg	1170
		Belagerte Demin uund Stettin gegen Kasimir I.	1180
		Gründete Kloster Arendsee	1183
		Stiftete erste Kloster Lehnin der Mark	1180
		Taufgeschenk war die Zauche	1126
Otto Normalverbraucher	De	erlernt Autofahren	1960
		erlernt eine Programmiersprache	2020
Otto von Bismarck	Preussen(De)	Emser Depesche	1870
		Erster Reichskanzler von Deutschland	1871
		mit Zuckerbrot und Peitsche-Politik	1871
Otto von Guericke	MD (De)	Erfinder der Luftpumpe	1649
		Halbkugelversuch vor dem Kaiser	1654
Otto Waalkes	Friesland De	Otto - Der Ausserfriesische	1990
		Otto - Der Film	1987

Wenn ich in diesem Beispiel alle Strupel entferne, die Ungarn enthalten, entferne ich auch zufällig das eine Strupel, das König enthält. D.h., wenn man oben König weglässt, ergibt sich das gleiche Ergebnis. Somit erhält man alle 9 Strupel der Ausgangstabelle `ottos.tab`, wenn man das letzte Ergebnis mit dem ersten zusammenfasst.

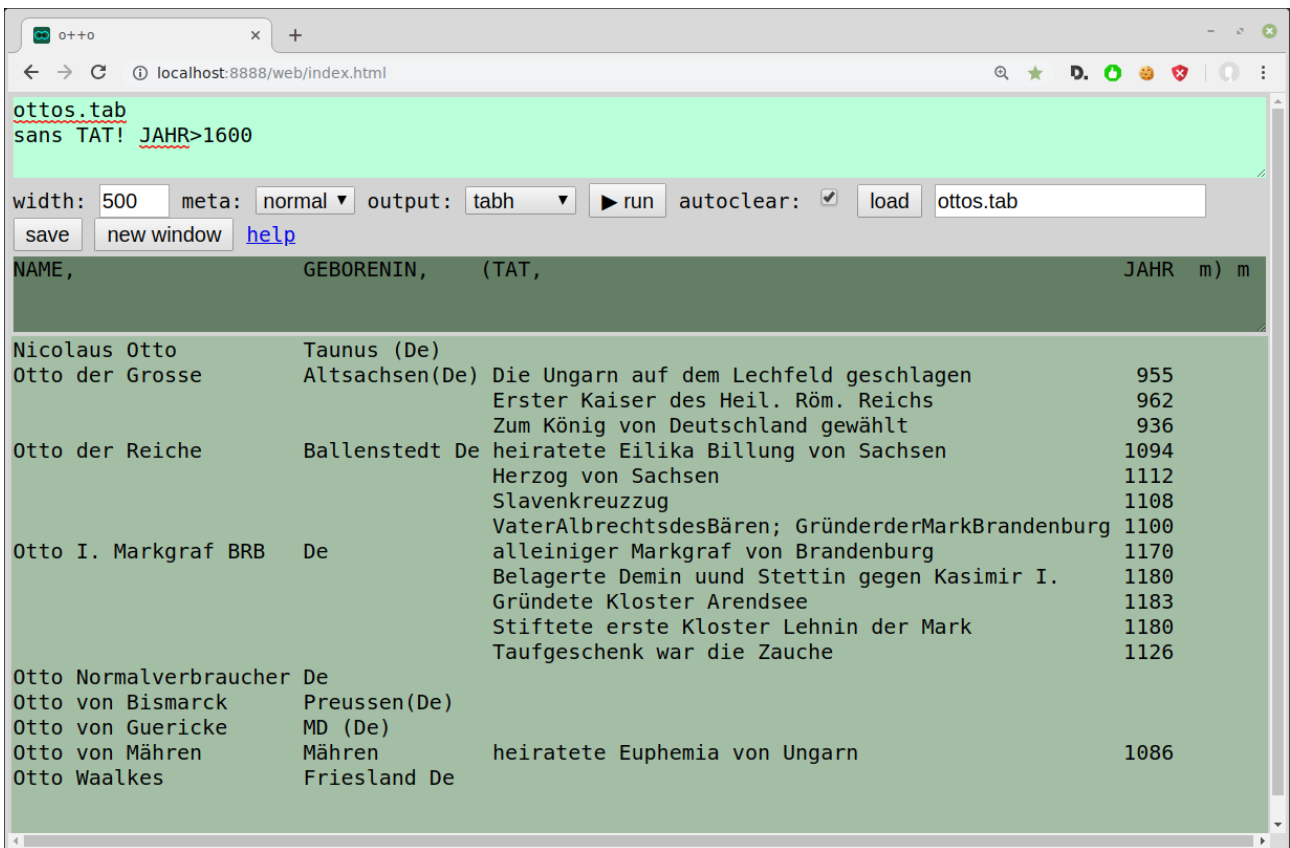
Man kann auch selektieren ohne dass sich die Anzahl der Strupel ändert, indem man lediglich (TAT, JAHR)-Subtupel selektiert:

```
ottos.tab
avec TAT! Ungarn
```



In analoger Weise betrachten wir das Programm:

ottos.tab
sans TAT! JAHR>1600



Man kann auch in beiden Kollektionen selektieren:

The screenshot shows a web browser window with the URL `localhost:8888/web/index.html`. The search query is `ottos.tab sans NAME TAT! JAHR>1600`. The results are displayed in a table with the following columns: NAME, GEBORENIN, (TAT, JAHR m) m.

NAME,	GEBORENIN,	(TAT,	JAHR m) m
Otto der Grosse	Altsachsen(De)	Die Ungarn auf dem Lechfeld geschlagen	955
		Erster Kaiser des Heil. Röm. Reichs	962
		Zum König von Deutschland gewählt	936
Otto der Reiche	Ballenstedt De	heiratete Eilika Billung von Sachsen	1094
		Herzog von Sachsen	1112
		Slavenkreuzzug	1108
		VaterAlbrechtsdesBären; GründerderMarkBrandenburg	1100
Otto I. Markgraf BRB De		alleiniger Markgraf von Brandenburg	1170
		Belagerte Demin uund Stettin gegen Kasimir I.	1180
		Gründete Kloster Arendsee	1183
		Stiftete erste Kloster Lehnin der Mark	1180
		Taufgeschenk war die Zauche	1126
Otto von Mähren	Mähren	heiratete Euphemia von Ungarn	1086

Diese Selektion ist gleichwertig zur Selektion
avec `JAHR>1600`

Man selektiert hierbei in allen Kollektionen, die JAHR enthalten. Das ist wieder sowohl
(TAT, JAHR m) als auch (NAME, GEBORENIN, (TAT, JAHR m) m).

Man erkennt, dass man durch ein Wort oder ein Zeichen mehr oder weniger völlig andere Anfragen
mit völlig anderen Ergebnissen spezifizieren kann. Das ist die Ursache dafür, dass o++o sehr
ausdrucksstark ist und die Programme trotzdem sehr kompakt sind. Da in anderen Papieren bereits
sehr viele Beispiele mit Selektionen enthalten sind, wird an dieser Stelle abgebrochen.