

Volltextsuche mit MS SQL-Server

Systemvoraussetzungen: Office Manager Enterprise 10.0 oder ab 9.0.4.391 mit gesonderter Aktivierung.

MS SQL-Server Vollversion ab 2000 (nicht MSDE oder SQL-Server Express)

Sie benötigen Administrator-Rechte für die Installation und Einrichtung!

Hinweise zur allgemeinen Einrichtung des MS SQL-Servers für Office Manager finden Sie unter

http://www.officemanager.de/support/omcs_mssql.pdf

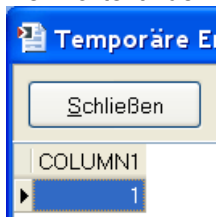
Letzte Änderung 08.05.2009

A. Anmelden

1. Melden Sie sich unter Windows bitte als Administrator an und stellen Sie sicher, dass Sie administrative Berechtigungen auf dem SQL-Server besitzen.
2. Wählen Sie im Office Manager den Menübefehl **Ansicht | Optionen | Administrator**. Öffnen Sie das Archiv, für welches die SQL-Server-Volltextsuche aktiviert werden soll.

B. Checkliste

1. Systemvoraussetzungen: Office Manager Enterprise 10.0 oder ab 9.0.4.391 mit gesonderter Aktivierung, MS SQL-Server Vollversion ab Version 2000 (nicht MSDE oder SQL-Server Express).
2. Wählen Sie im Office Manager-Menü **Suchen** bitte den Befehl **SQL-Abfrage**. Geben Sie `SELECT fulltextserviceproperty('IsFulltextInstalled')` ein und wählen Sie den Menübefehl **Abfrage | Temporäre Ergebnismenge verwenden**. Der Volltextindex ist auf dem Server installiert, wenn eine 1 zurückgegeben wird:



3. Das geöffnete Archiv muss eine Datenbank auf einem MS SQL-Server sein, die Endung der Archivdatei lautet .omcs

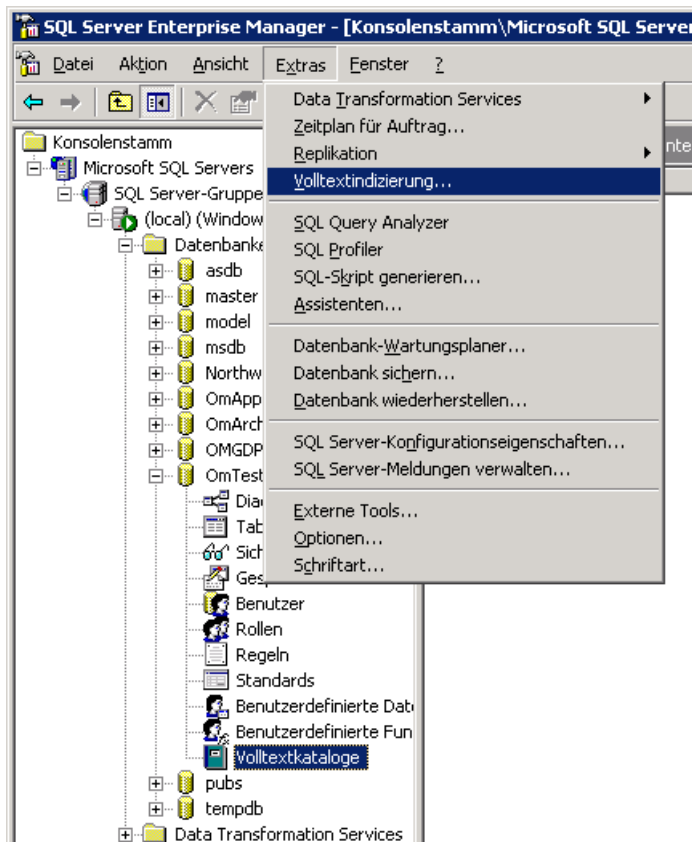
C. Archiv im Office Manager erweitern

Der SQL-Server indexiert die Felder der Datenbank, aber nicht die verlinkte Dokumentdatei. Für eine Suche im Dokumententext muss dieser in die Datenbank kopiert werden.

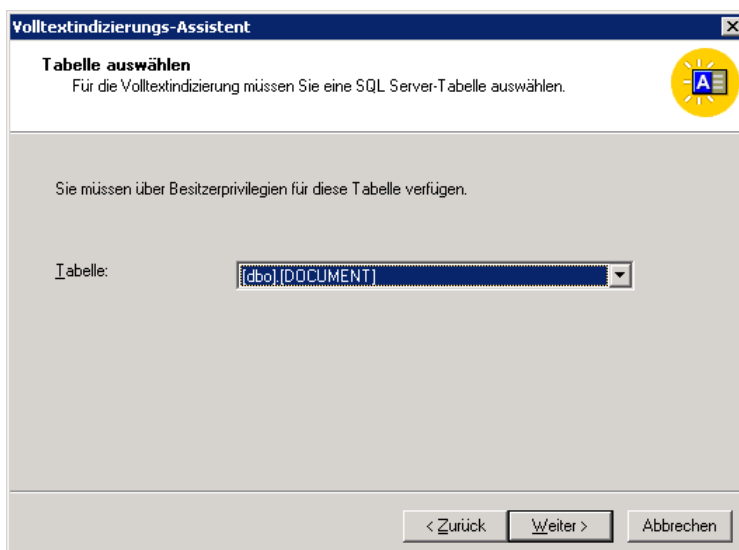
1. Wählen Sie den Menübefehl **Datei | Datenbank | Felder definieren** und klicken Sie auf den Schalter **Neu**. Geben Sie bitte folgende Eigenschaften ein:
Interner Name: CONTENTTEXT
Datentyp: Text (Memo)
Bezeichnung: z.B. Volltext
Sichtbar: für niemanden
Berechtigung zum Bearbeiten: für niemanden
2. Klicken Sie bitte auf **Übernehmen** und anschließend auf **Schließen**.
3. Wählen Sie den Menübefehl **Datei | Information**. Auf Seite **Volltextindex** muss jetzt "SQL Server Volltextindex" stehen.

D. Indexierung des Servers aktivieren

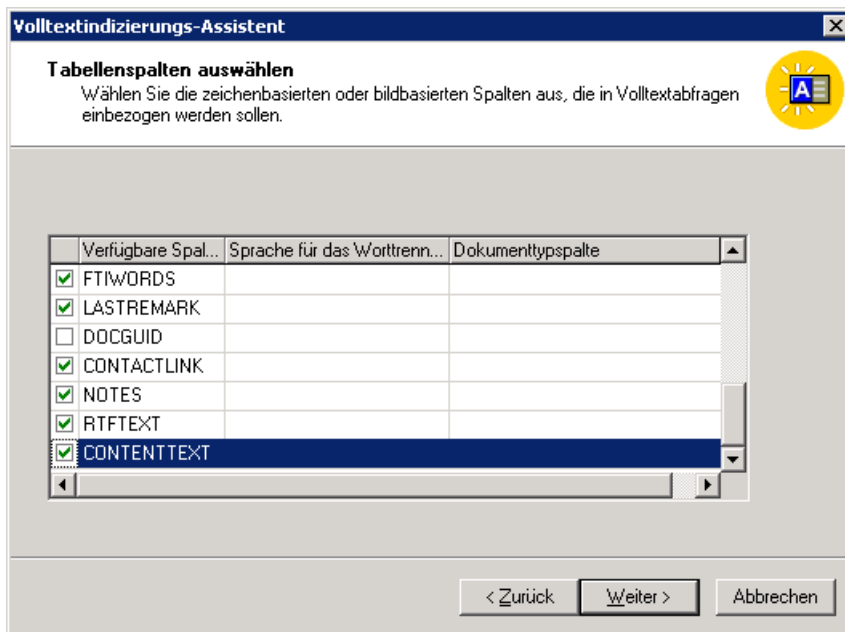
1. Starten Sie bitte die SQL-Server Benutzeroberfläche Management Studio bzw. Enterprise Manager.
2. Öffnen Sie die Datenbank im Baum und markieren Sie Volltextkataloge. Starten Sie mit **Extras | Volltextindexierung** den Assistenten:



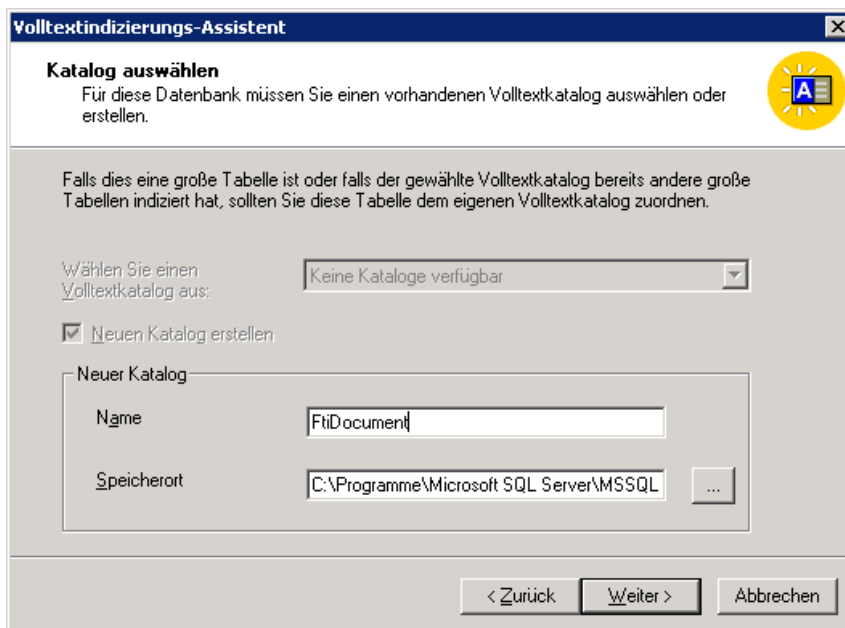
3. Wählen Sie bitte die Tabelle DOCUMENT und im nächsten Fenster den Index "PK__DOCUMENT__...":



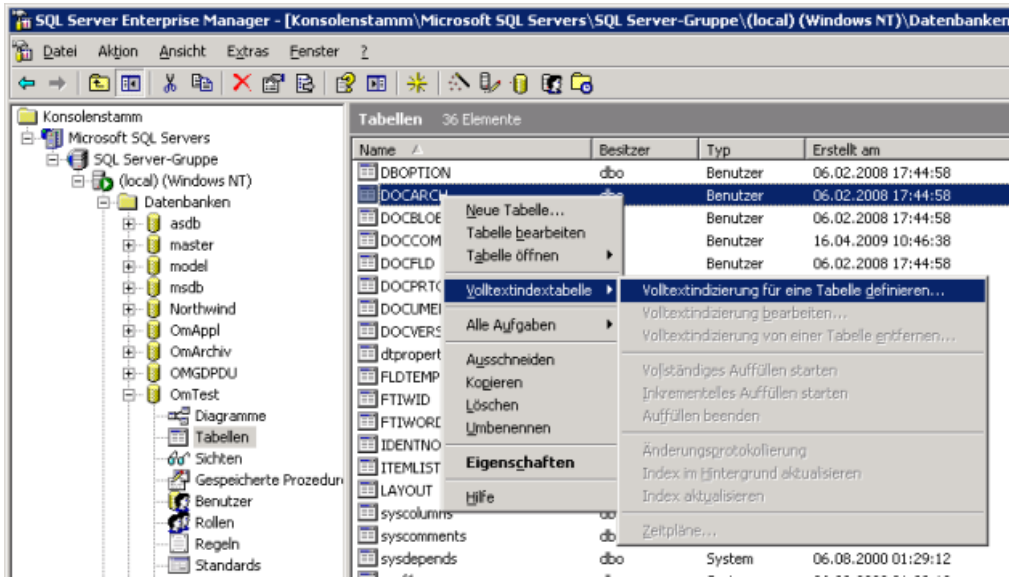
4. Spalten für die Indexierung wählen. Folgende Spalten sollten nicht markiert werden, weil sie keine relevanten Daten enthalten: FAVORITE, DOCTYPE, DOCGUID, MAILATTACHICO, alle benutzerdefinierten Felder vom Datentyp "Logisch (ja/nein)" und die Felder, die Sie nicht im Volltextindex wünschen. Markieren Sie bitte alle anderen Tabellenspalten:



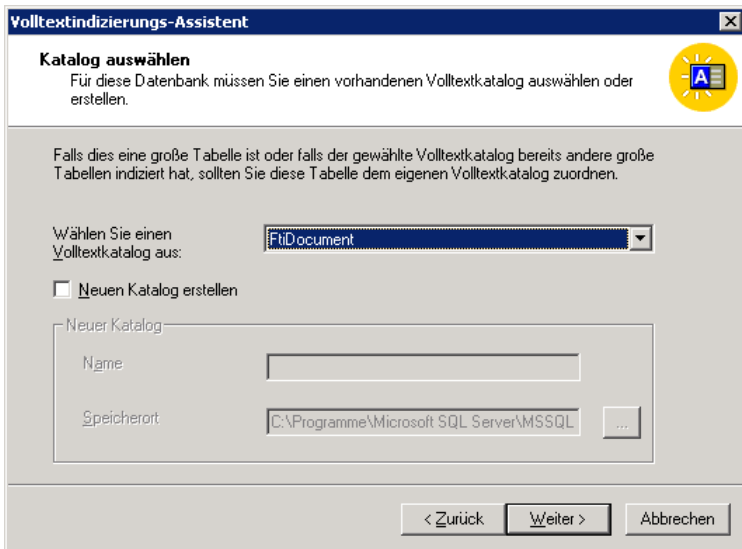
5. Erstellen Sie einen neuen Katalog mit Namen "FtiDocument":



6. Markieren Sie im Baum die Tabelle DOCARCH, klicken Sie diese mit der rechten Maustaste an und wählen Sie im lokalen Menü den Befehl **Volltextindextabelle | Volltextindizierung für diese Tabelle definieren**:



7. Wählen Sie bitte den Index "PK__DOCARCH___...", markieren Sie die gewünschten Felder und den zuvor erstellten Volltextkatalog:



Klicken Sie bitte mehrmals auf **Weiter** und anschließend auf **Fertigstellen**.



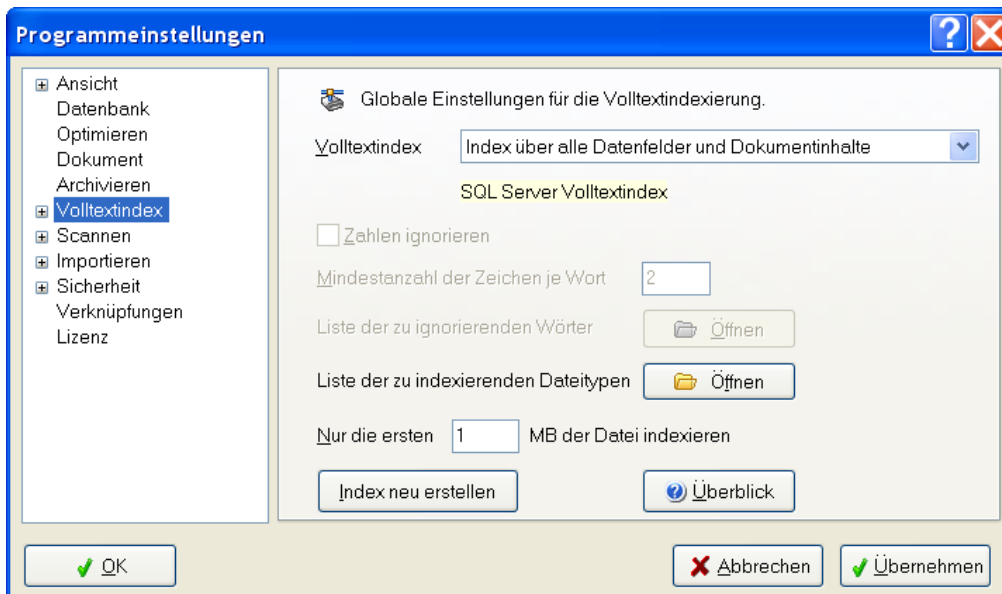
Führen Sie das Auffüllen jetzt noch nicht durch!

E. Dokumententext extrahieren

Diesen Schritt können Sie überspringen, wenn das Archiv noch keine Dokumente beinhaltet.

Alle Dokumente des Archivs müssen vom Office Manager gelesen und der enthaltene Text in das Datenfeld CONTENTTEXT kopiert werden. Dieser Schritt ist bei großen Archiven sehr zeitaufwendig und sollte beispielsweise an einem arbeitsfreien Wochenende durchgeführt werden.

1. Wählen Sie bitte im Office Manager-Menü **Ansicht | Optionen** den Befehl **Einstellungen** und wechseln Sie auf die Seite **Volltextindex**:

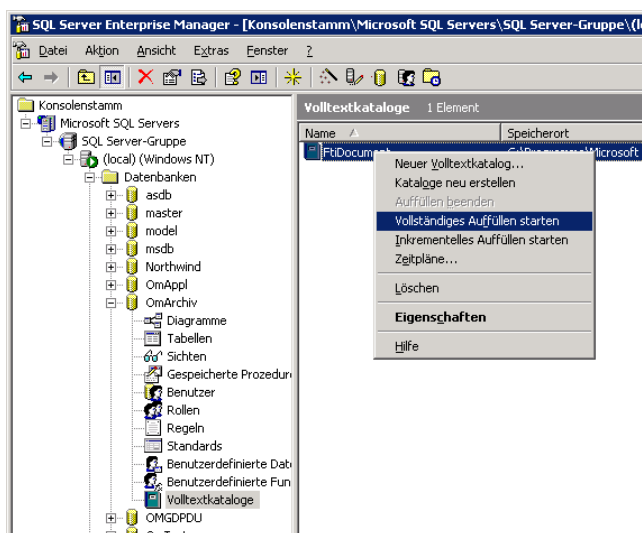


Wählen Sie für **Volltextindex** die Option "Index über alle Datenfelder und Dokumentinhalte".

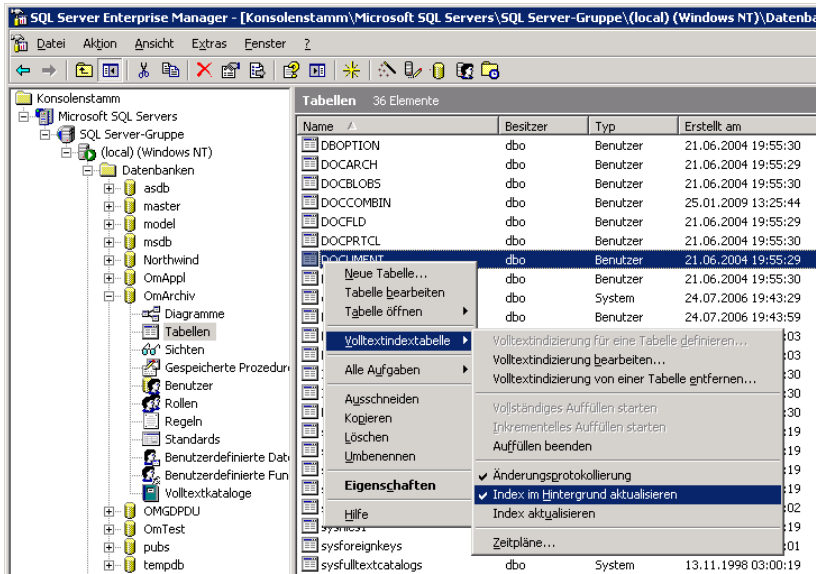
2. Bei **Liste der zu indexierenden Dateitypen** legen Sie fest, welche Dokumentdateien vom Office Manager konvertiert und in das Feld CONTENTTEXT kopiert werden. Dateien mit folgender Endung können konvertiert werden: doc, docm, docx, dotm, dotx, eml, htm, html, jpg, jpeg, mht, mhtml, msg, odf, odp, ods, odt, pdf, ppsx, pptx, rtf, txt, tif, tiff, sxc, sxd, sxi, sxw, xls, xlsx, wpd, zip und alle Dateien im Textformat.
3. Klicken Sie bitte auf den Schalter **Index neu erstellen**. Die Volltexttabellen FTIWID und FTIWORD des Archivs werden gelöscht, die Dokumente konvertiert und der Text in CONTENTTEXT kopiert.

F. Serverindexierung starten

1. Markieren Sie in der Serversoftware bitte den Eintrag **Volltextkataloge** unterhalb der Datenbank, klicken Sie mit der rechten Maustaste auf FtIDocument und wählen Sie **Vollständiges Auffüllen starten**:



2. Sobald die Funktion abgeschlossen ist, aktivieren Sie bitte die automatische Index-Aktualisierung. Klicken Sie dazu bitte mit der rechten Maustaste auf die Tabelle DOCUMENT und markieren Sie im lokalen Menü den Eintrag **Volltextindextabelle | Änderungsprotokollierung**. Klicken Sie erneut mit der rechten Maustaste auf DOCUMENT und markieren Sie **Volltextindextabelle | Index im Hintergrund aktualisieren**:



Wiederholen Sie dies anschließend für die Tabelle DOCARCH.

Sie können alternativ auch andere Funktionen (z.B. Zeitpläne) für die Index-Aktualisierung verwenden. Lesen Sie hierzu bitte die SQL-Server-Dokumentation.

G. Informationen zum Volltextindex des MS SQL-Servers

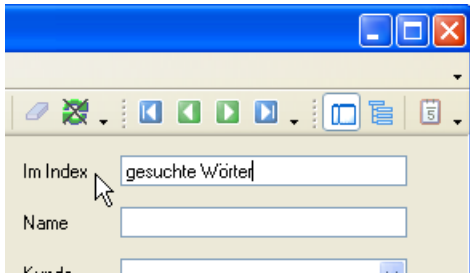
Dieses Dokument finden Sie unter: <http://www.officemanager.de/support/omcs-mssql-volltext.pdf>

Archive auf MS SQL: http://www.officemanager.de/support/omcs_mssql.pdf

SQL-Server-Dokumentation: <http://technet.microsoft.com/de-de/library/ms142571.aspx>

H. Volltextsuche mit dem Office Manager

Die Suche können Sie im primären Suchfeld **Im Index** oder im Abfragedialog (Menü **Suchen | Volltextindex**) eingeben.



1. Hinweise

- Die Volltextsuche berücksichtigt keine Groß-/ Kleinschreibung.
- Operatoren (AND, OR, ...) müssen groß geschrieben werden.
- Gesuchte Wörter sollten klein geschrieben werden, um eine Verwechslung mit Operatoren zu vermeiden.
- Die Standard-Verknüpfung ist AND. Wenn die Suche mehrere Wörter enthält, dann werden nur Dokumente gefunden, die alle Wörter enthalten.
- Wenn keine Operatoren oder Anführungszeichen verwendet werden, dann werden auch ähnliche Wörter gefunden. *Die Abfrage erfolgt dann intern mit FREETEXT.*
- Wenn ein Operator, * oder Anführungszeichen verwendet werden, dann muss die Suche nach ähnlichen Wörter explizit angegeben werden (~, ~~, INFL, SYN). *Die Abfrage erfolgt intern mit CONTAINS.*

2. Einfache Suche nach einem oder mehreren Wörtern

Geben Sie die Wörter einfach mit Leerzeichen getrennt ein, z.B. `wort1 wort2 wort3`. Es werden alle Dokumente gefunden, die alle gesuchten Wörter oder ähnliche Wörter beinhalten.

3. Suche nach einem Ausdruck (einer Phrase)

Geben Sie den gesuchten Ausdruck bitte in doppelten Anführungszeichen ein, z.B. `"office manager dms"`. Es werden nur Dokumente gefunden, die genau diesen Wortlaut beinhalten.

4. Erweiterte Suche mit Operatoren

Beachten Sie bitte, dass die Operatoren GROSS geschrieben werden müssen.

- **AND**: sucht Dokumente, die alle Wörter beinhalten, z.B. `wort1 AND wort2`.
Alternative Schreibweise: `wort1 & wort2`.
- **OR**: sucht Dokumente, die mindestens ein Wort beinhalten, z.B.: `wort1 OR wort2`.
Alternative Schreibweise: `wort1 | wort2`.
- **AND NOT**: sucht Dokumente die das Wort 1 beinhalten. In dem Datenfeld, in welchem Wort 1 gefunden wurde, darf Wort 2 aber nicht vorkommen,
z.B.: `wort1 AND NOT wort2`.
Alternative Schreibweise: `wort1 &! wort2`.
- **NEAR**: Wort 2 muss im selben Datenfeld und in der Nähe von Wort 1 stehen: `wort1 NEAR wort2`.
- **INFL**: für Inflectional sucht ähnliche Wörter mit Hilfe der grammatikalischen Beugung, z.B. `INFL gehen`:
findet Dokumente mit dem Wort gehen, ging oder geht.
Alternative Schreibweise: `~gehen`.

- **SYN** für Synonym sucht ähnliche Wörter mit Hilfe eines Thesaurus, z.B. `SYN auto` findet Dokumente mit dem Wort auto, automobil oder pkw. *Das Suchergebnis ist abhängig vom Thesaurus des SQL-Servers.*
Alternative Schreibweise: `~~auto`.
- *****: sucht Wörter oder Phrasen, die mit dem eingegebenen Text beginnen, z.B. `offi*` findet Dokumente mit dem Wort office, offiziell u.a.
- **COL** begrenzt die folgende Suche auf ein Datenbankfeld, z.B. `COL NAME wort` findet nur Dokumente, in denen das Wort im Dokumentattribut (Datenbankfeld) NAME vorkommt. Das Datenfeld für den Dokumenteninhalte heißt CONTENTTEXT, `COL CONTENTTEXT wort` sucht also nur im Dokument selbst.
`COL * wort` sucht wieder in allen Datenbankfeldern.
Alternative Schreibweise für COL CONTENTTEXT wort: `DOCUMENT wort`.

5. Expertensuche mit SQL-Anweisung

Wenn der eingegebene Suchtext mit `CONTAINS (`, `FREETEXT (`, `CONTAINSTABLE (` oder `FREETEXTTABLE (` beginnt, dann wird der gesamte Suchtext direkt im WHERE-Teil der SQL-Abfrage übergeben.

Diese Art der Suche ist für administrative Kontrollabfrage vorgesehen und für den Anwender nicht geeignet.

6. Beispiele

`raben fliegen`: findet z.B. Dokumente mit dem Text "Raben können fliegen", "der Rabe fliegt nach Berlin" oder "die Raben flogen heute schnell".

`"raben können fliegen"`: findet nur Dokumente mit dem genauen Ausdruck "Raben können fliegen".

`"raben können fliegen" COL APPLNAME acrobat`: findet nur Dokumente, die den genauen Ausdruck "Raben können fliegen" enthalten und das Wort "Acrobat" im Datenfeld APPLNAME (Dokumenttyp) beinhalten.

`raben OR fliegen`: Dokumente, in denen das Wort "Raben" oder "fliegen" vorkommt.

`~raben OR ~fliegen` oder `INFL raben OR INFL fliegen`: Dokumente, in denen eines der folgenden Wörter vorkommt: "Rabe", "Raben", "fliegt", "flog", "fliegen" u.ä.

`~~rabe` oder `SYN rabe`: abhängig vom verwendeten Thesaurus z.B. "Rabe", "Vogel", oder "Krähe".

`COL NAME raben`: das Wort "Raben" muss im Dokumentattribut NAME enthalten sein.

`COL NAME raben COL * fliegen`: Das Wort "Raben" muss im Dokumentattribut NAME und das Wort "fliegen" in einem beliebigen Attribut oder in der Dokumentdatei enthalten sein.

`COL CONTENTTEXT "raben können fliegen"` oder `DOCUMENT "raben können fliegen"`: der Ausdruck "Raben können fliegen" muss in der Dokumentdatei vorkommen.

`raben NEAR fliegen`: beide Wörter "Raben" und "fliegen" müssen im selben Dokumentattribut (oder der Datei) enthalten sein und nahe beieinander stehen.

`rabe* OR krähe*`: findet Dokumente, die ein Wort enthalten, das mit Rabe oder Krähe beginnt.

`pinguine &! fliegen`: das Wort "Pinguine" muss vorhanden sein und das Wort "fliegen" darf nicht im selben Attribut vorkommen.