



Lösungen für mittelständische Unternehmen

Microsoft SQL Server: File Tables



Nutzung für Dokumentenmanagementsysteme

Copyright © 2016 ITSM OHG



ITSM OHG

Elisabeth-Selbert-Str. 19a

40764 Langenfeld / Rheinland

www.itsm-erp.de

info@itsm-erp.de

Fon +49 (0) 2173 10648-0

Fax +49 (0) 2173 10648-48



Management Summary

Der Nutzen von Dokumentenmanagementsystemen (kurz DMS) ist inzwischen weitgehend unbestritten. Wer kennt nicht die Situation, dass es von einem Dokument mehrere Versionen gibt und man sich fragt, welches die aktuelle ist?

Die Frage ist nur, bringt es eine weitere Insellösung, die zwar sehr gut Dokumente verwalten kann, aber mehr auch nicht? Wir sind der Meinung, dass ein DMS als Insellösung zwar besser ist als nichts, aber seine wirklichen Fähigkeiten nur dann ausspielen kann, wenn es integriert ist in eine umfassende ERP/CRM-Lösung, die Dokumente mit Stammdaten, Personen, Projekten oder anderen Elementen verknüpft.

Ist es mit vertretbarem Aufwand möglich, selber eine entsprechende Lösung zu schaffen, die Schluss mit dem häufig sehr zeitraubenden Suchen nach Dokumenten macht? Die Antwort ist „Im Prinzip ja, ...“.

Die neueren Versionen des SQL Server bieten mit den File Tables einen technisch gut beherrschbaren Ansatz, Volltextsuchen in Dokumenten unterschiedlichster Arten (einschließlich gescannter PDFs) zu realisieren.

Die folgenden Seiten zeigen detailliert die Vorgehensweise zur Einrichtung von File Tables als Basis für eine integrierte Lösung einer Volltextsuche auf, gehen jedoch über den technischen Ansatz nicht hinaus. Das würde den Rahmen dieses Dokuments eindeutig sprengen.

Es handelt sich um eine technisch orientierte Dokumentation, die sich an IT-Fachleute wendet, die bereits grundlegende Erfahrungen mit dem Microsoft SQL Server haben.



Allgemeine Hinweise

Das Dokument beschreibt die Einrichtung von File Tables mit Full-Text Index als Voraussetzung für eine Volltextsuche im Dokumentenmanagementsystem (DMS). Die Anleitungen und Screenshots beziehen sich auf den SQL Server 2016.

Das Dokument kann als Tutorial verwendet werden. In den Skripten wird eine Datenbank mit zwei File Table Tabellen angelegt. Um die Korrelation zwischen Objektnamen in der Datenbank und im Dateisystem zu veranschaulichen, wurde eine zweite vergleichbare Datenbank mit einer File Table angelegt. Dafür werden keine Skripte gezeigt. Da das Tutorial grundlegende Kenntnisse des SQL Server voraussetzt, sollte dies für den Datenbankadministrator kein Problem sein.

ITSM empfiehlt, zunächst die Beispiele in einer Testumgebung auszuführen, bevor Sie an Ihre Produktiv-Umgebung herangehen. ITSM übernimmt keine Gewähr dafür, dass diese Anleitungen fehlerfrei sind und in jeder Betriebsumgebung so funktionieren.

Bitte sorgen Sie immer für vollständige Backups, bevor Sie mit neuen Funktionen experimentieren.

Anleitung im Detail

Die folgende Tabelle zeigt die notwendigen und optionalen Schritte zur Einrichtung:

Step	Aktion	Nur für Volltextsuche	Bemerkung
1	Full-Text Feature installieren	x	SQL Server Installation Center
2	FileStream aktivieren		SQL Server Configuration Manager
3	Zugriffslevel konfigurieren		SQL Server Management Studio (SSMS)
4	Datenbank anlegen		SSMS
5	FileGroup anlegen		SSMS
6	File in FileGroup anlegen		SSMS
7	Directory Name festlegen		SSMS
8	File Table Tabelle anlegen		SSMS
9	Full-Text Catalog einrichten	x	SSMS
10	Dokumente aufnehmen		Explorer
11	Full-Text Index definieren	x	SSMS

Schritt 1 – Full-Text Feature installieren

Dieser Schritt ist nur dann erforderlich, wenn Sie auch eine Volltextsuche einrichten möchten. Für das Erstellen von File Tables muss das Feature nicht installiert sein.

Bei der Installation der SQL Server-Instanz muss die Option Full-Text mit ausgewählt werden. Dieses Feature kann jedoch auch jederzeit über das SQL Server Installation Center nachinstalliert werden.

Achtung: Das Feature steht in der SQL Server Express-Edition nicht zur Verfügung.

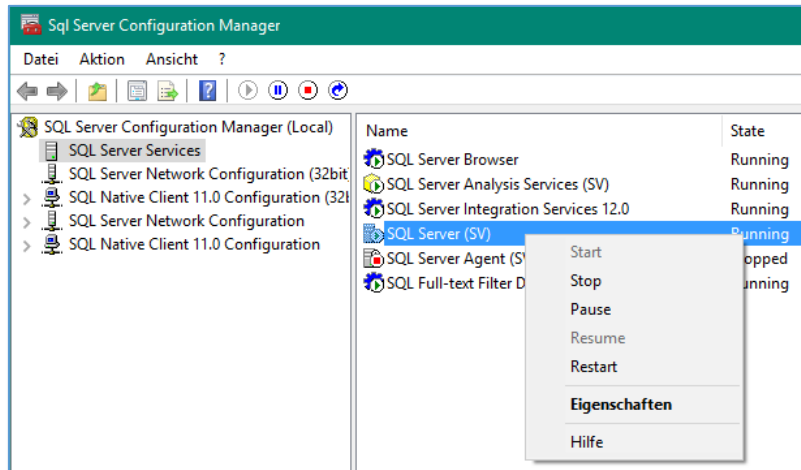
Schritt 2 – FileStream aktivieren

Der Dienst zum Starten der SQL Server Instanz muss entsprechend konfiguriert werden.

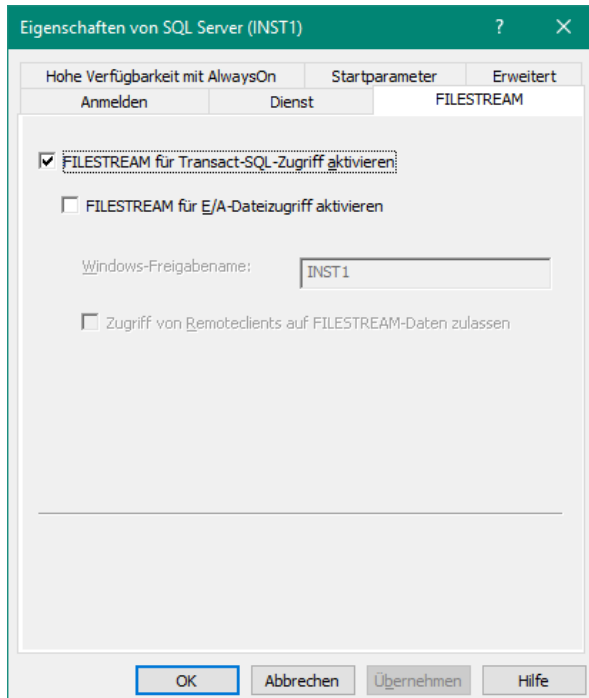


Starten Sie den SQL Server Configuration Manager.

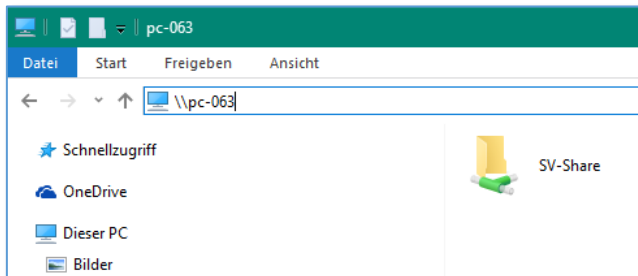
Rufen Sie dazu die Eigenschaften des SQL Server Dienstes auf:



Aktivieren Sie die FileStream-Funktion. Sie müssen einen Namen für ein Windows-Share vergeben, die global für die gesamte SQL Server-Instanz gilt.

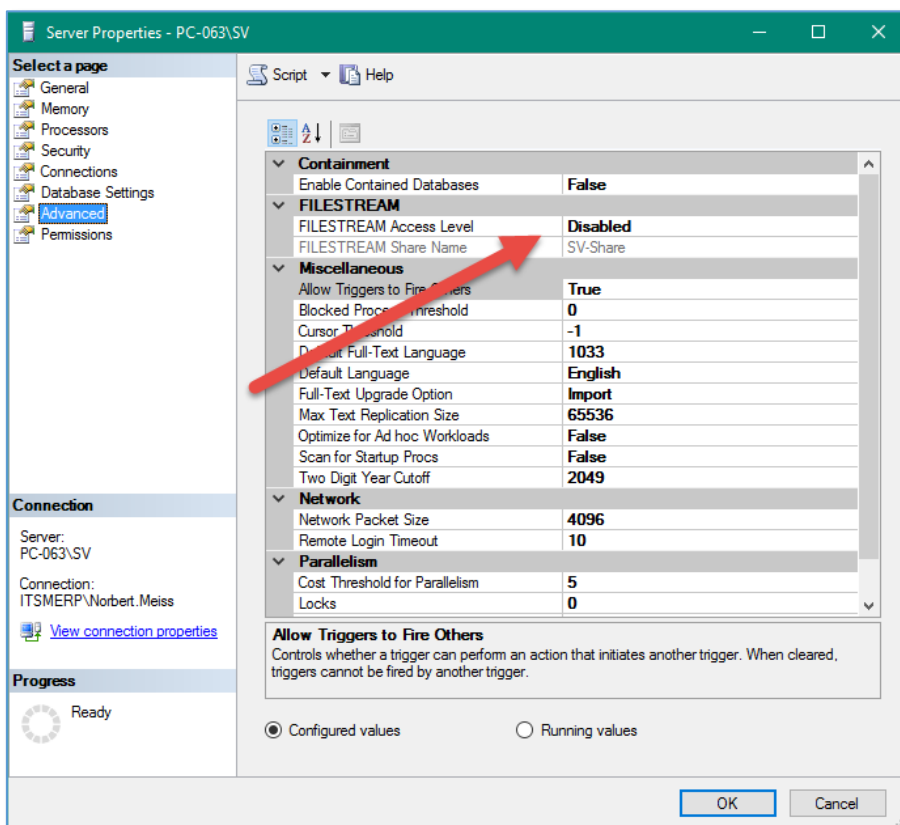


Dieses Share steht ganz normal im Dateisystem zur Verfügung. Es handelt sich jedoch trotzdem nicht um ein übliches Windows-Share, für das Sie zum Beispiel gezielt Berechtigungen vergeben könnten.



Über die Folderstruktur der Share können Dokumente dem SQL Server „bekannt“ gemacht werden. Dies wird später beispielhaft gezeigt.

In den Eigenschaften des SQL Server ist das Share zu sehen, allerdings noch mit dem Access Level „Disabled“:



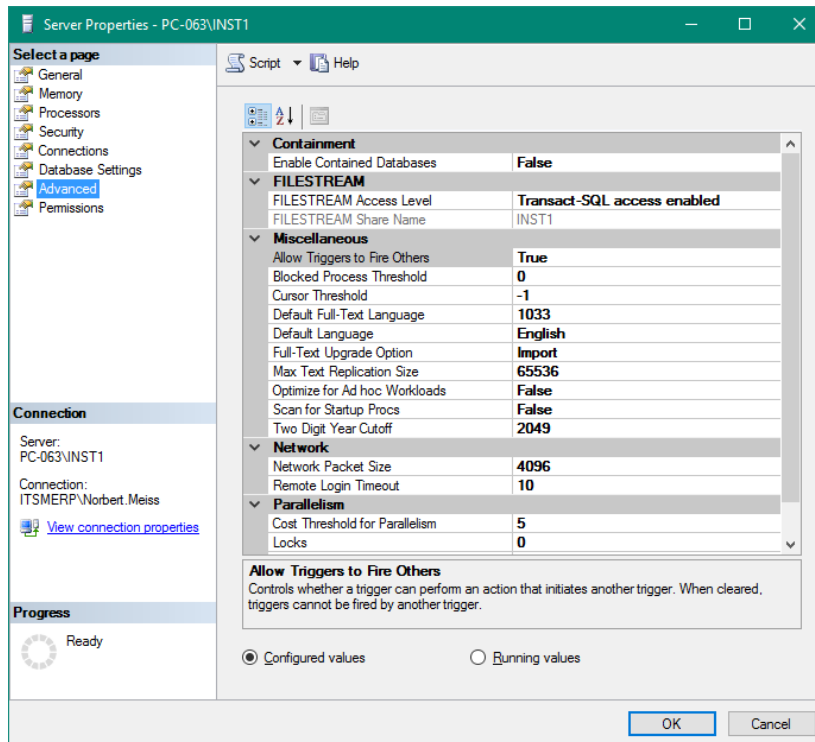
Schritt 3 – Zugriffslevel konfigurieren

Führen Sie im SQL Server Management Studio (SSMS) das folgende Skript aus:

```
EXEC sp_configure filestream_access_level, 2;  
GO  
RECONFIGURE  
GO
```



Damit ist der uneingeschränkte Zugriff aktiviert:



Schritt 4 – Datenbank anlegen

Führen Sie im SSMS das folgende Skript aus:

```
USE master;
GO

CREATE DATABASE EX1 ON
( NAME = EX1_dat,
  FILENAME = 'C:\Program Files\Microsoft SQL Server\MSSQL12.SV\MSSQL\DATA\EX1_dat.mdf' )
LOG ON
( NAME = EX1_log,
  FILENAME = 'C:\Program Files\Microsoft SQL Server\MSSQL12.SV\MSSQL\DATA\EX1_log.ldf' )
GO
```

Schritt 5 – FileGroup anlegen

File Table Tabellen können nur in einer speziellen FileGroup angelegt werden, also nicht in der FileGroup Primary. Führen Sie im SSMS das folgende Skript aus:

```
ALTER DATABASE EX1
ADD FILEGROUP EX1_fg
CONTAINS FILESTREAM;
GO
```

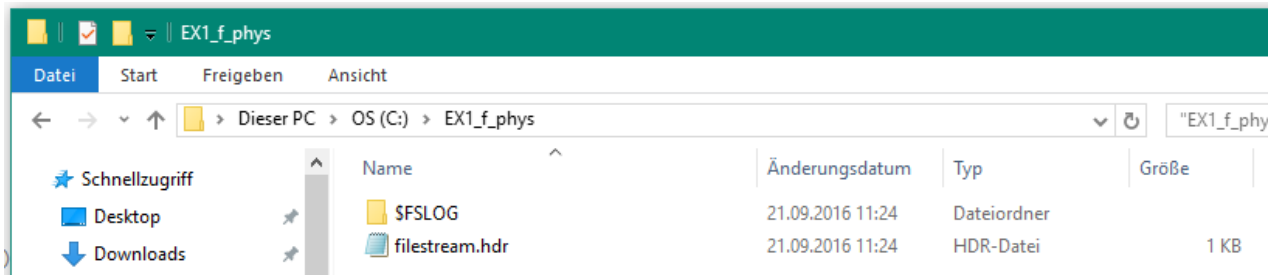
Schritt 6 – File in FileGroup anlegen

Wie bei den anderen Files auch, unterscheidet der SQL Server auch in einer Filestream FileGroup zwischen logischen und physischen Dateinamen. Führen Sie im SSMS das folgende Skript aus:

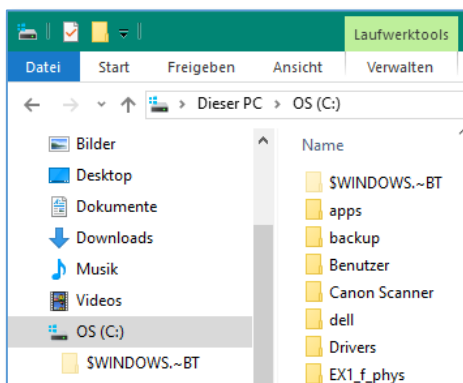
```
ALTER DATABASE EX1
```



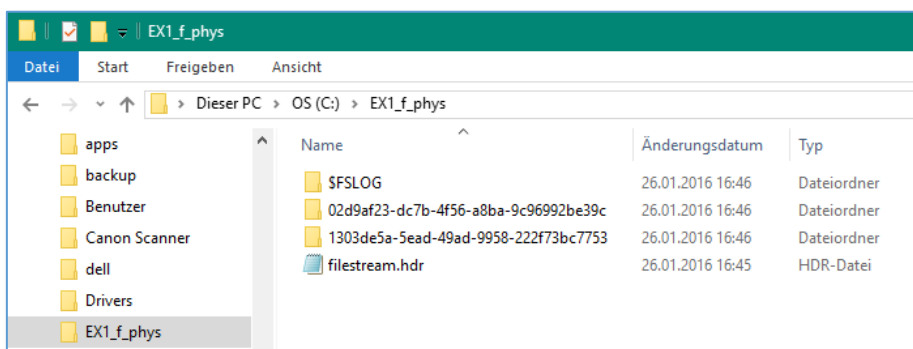
```
ADD FILE (
NAME = EX1_f_log,
FILENAME = 'C:\EX1_f_phys')
TO FILEGROUP EX1_fg;
GO
```



Die Bezeichnung des Parameters FILENAME ist etwas irreführend, denn tatsächlich wird im Dateisystem ein gleichnamiger Folder angelegt:



Die folgende Abbildung zeigt den anfänglichen Inhalt des Folders:



Schritt 7 – Directory Namen festlegen

Führen Sie im SSMS das folgende Skript aus:

```
ALTER DATABASE EX1 Set filestream (non_transacted_access = full,
Directory_Name = 'EX1_Dir')
GO
```



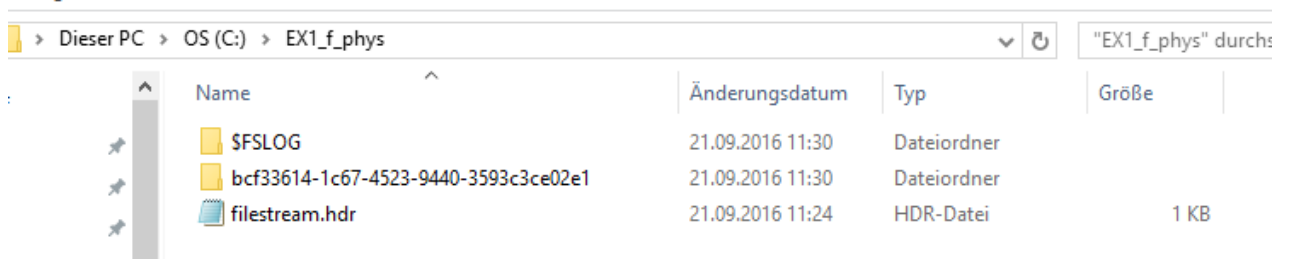
Schritt 8 – File Table Tabellen anlegen

Wechseln Sie jetzt im SSMS in die Beispieldatenbank EX1.

```
USE EX1  
GO
```

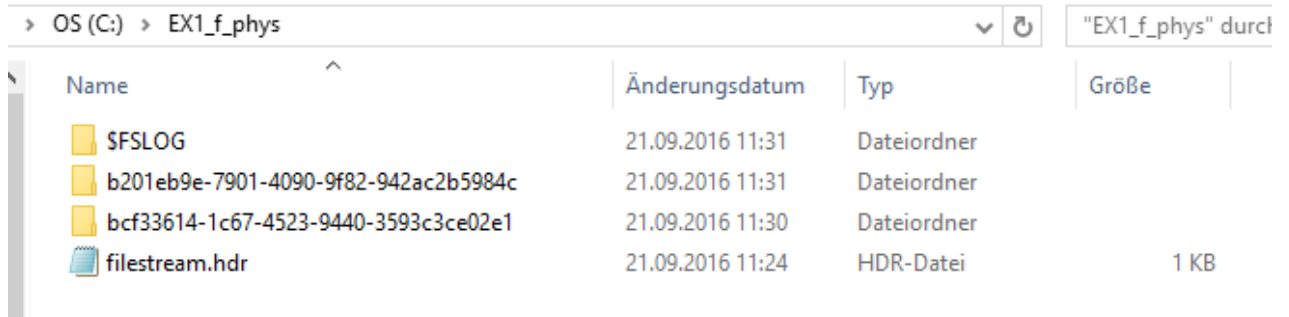
Führen Sie im SSMS das folgende Skript aus:

```
CREATE TABLE EX1_ft1 As FileTable FILESTREAM_ON EX1_fg  
GO
```



Name	Änderungsdatum	Typ	Größe
\$FSLOG	21.09.2016 11:30	Dateiordner	
bcf33614-1c67-4523-9440-3593c3ce02e1	21.09.2016 11:30	Dateiordner	
filestream.hdr	21.09.2016 11:24	HDR-Datei	1 KB

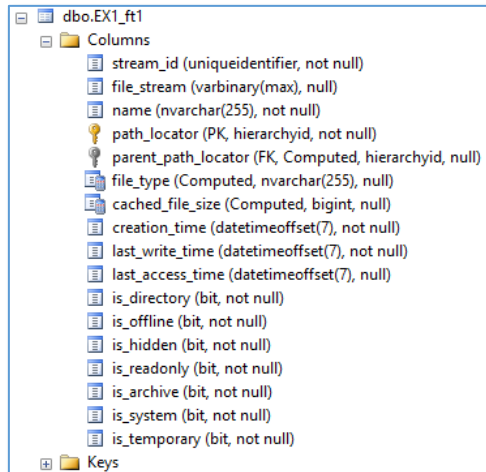
```
CREATE TABLE EX1_ft2 As FileTable FILESTREAM_ON EX1_fg  
GO
```



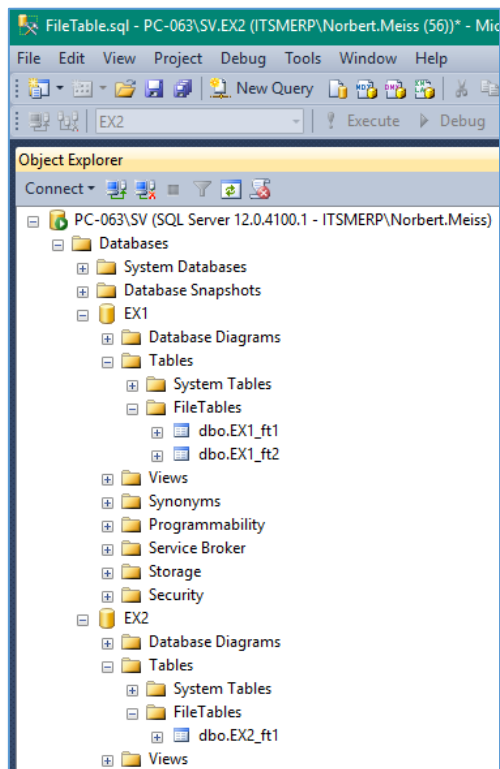
Name	Änderungsdatum	Typ	Größe
\$FSLOG	21.09.2016 11:31	Dateiordner	
b201eb9e-7901-4090-9f82-942ac2b5984c	21.09.2016 11:31	Dateiordner	
bcf33614-1c67-4523-9440-3593c3ce02e1	21.09.2016 11:30	Dateiordner	
filestream.hdr	21.09.2016 11:24	HDR-Datei	1 KB



Durch den Parameter As FileTable ist die Struktur der Tabelle vorgegeben:



Wenn Sie auch selber noch die eingangs erwähnte Datenbank EX2 angelegt haben, sollten Sie im SSMS Object Explorer jetzt folgende Struktur sehen:



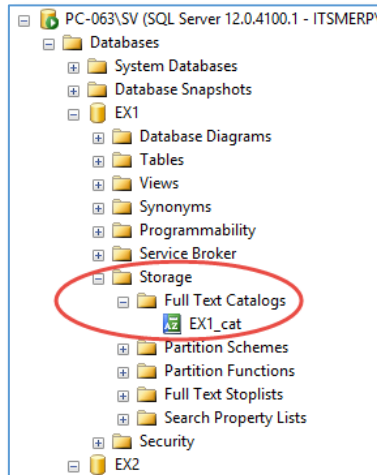


Schritt 9 – Full-Text Catalog einrichten

Führen Sie im SSMS das folgende Skript aus:

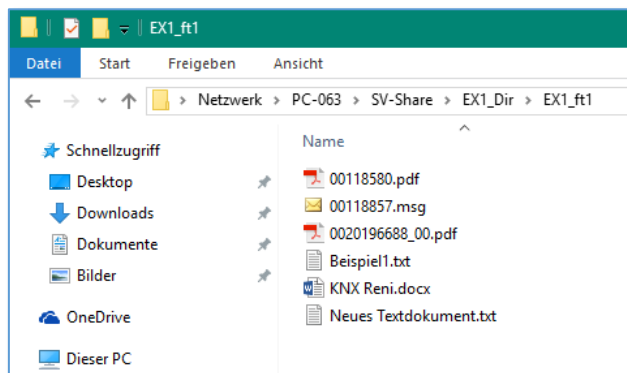
```
CREATE FullText Catalog EX1_cat
```

Sie finden den neuen Katalog im Knoten Storage | Full Text Catalogs:



Schritt 10 – Dokumente aufnehmen

Sie können jetzt die Dokumente, die Sie mit der neuen File Table Tabelle synchronisieren wollen, in der Share anlegen – entweder durch Neuerstellen oder Kopieren / Verschieben aus anderen Foldern. Wir kopieren jetzt einige Dateien in die Share \\pc-063\SV-Share\EX1_Dir\EX1_ft1:

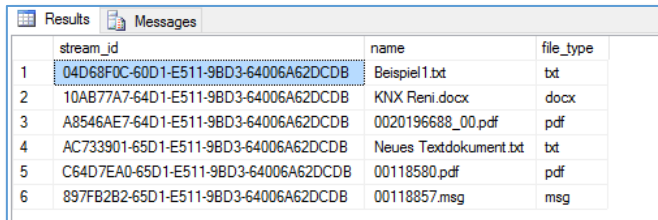


Wenn Sie sich jetzt die Records in der File Table EX1_ft1 ansehen, sehen Sie darin sofort die neuen Datensätze für diese Dateien.



Führen Sie dazu im SSMS das folgende Skript aus:

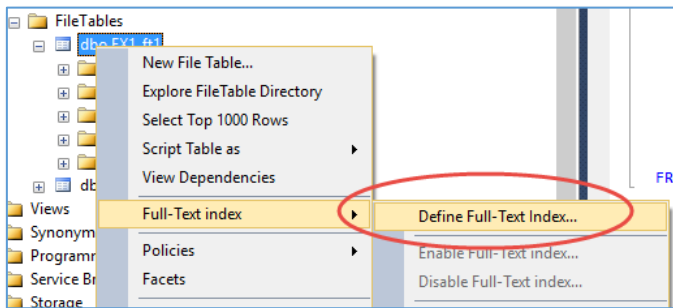
```
SELECT stream_id, name, file_type
FROM EX1_ft1
```



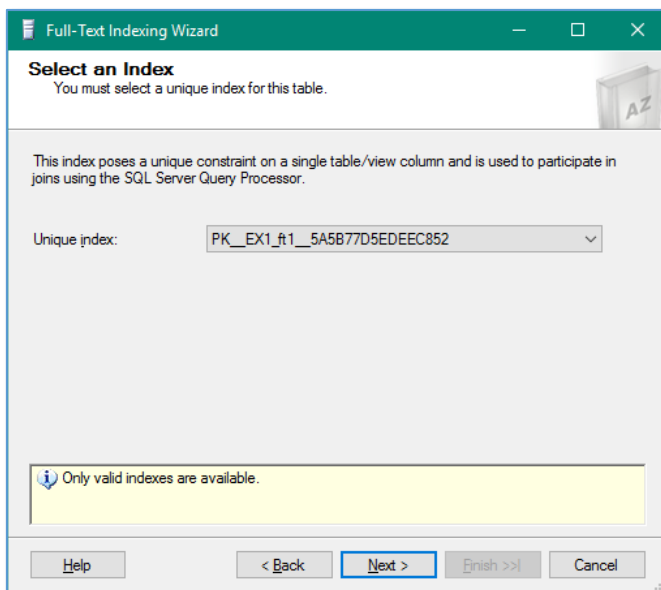
	stream_id	name	file_type
1	04D68FOC-60D1-E511-9BD3-64006A62DCDB	Beispiel1.txt	txt
2	10AB77A7-64D1-E511-9BD3-64006A62DCDB	KNX Reni.docx	docx
3	A8546AE7-64D1-E511-9BD3-64006A62DCDB	0020196688_00.pdf	pdf
4	AC733901-65D1-E511-9BD3-64006A62DCDB	Neues Textdokument.txt	txt
5	C64D7EA0-65D1-E511-9BD3-64006A62DCDB	00118580.pdf	pdf
6	897FB2B2-65D1-E511-9BD3-64006A62DCDB	00118857.msg	msg

Schritt 11 – Full-Text Index definieren

Im letzten Schritt müssen wir noch den Full-Text Index einrichten. Rufen Sie hier den Full-Text Indexing Wizard auf:

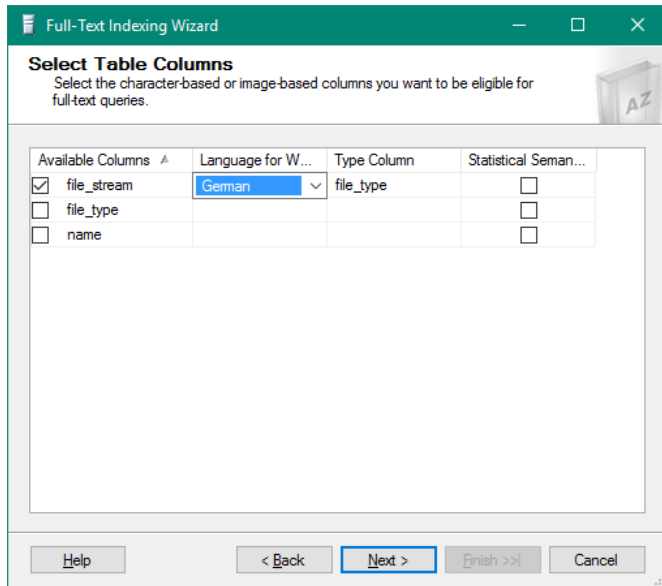


Die meisten Voreinstellungen im Wizard können Sie einfach mit Next bestätigen:





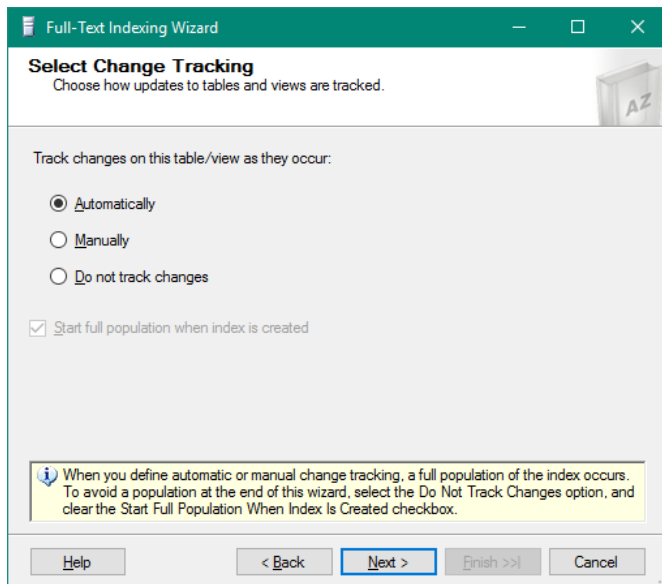
In diesem Schritt müssen Sie manuell eingreifen und angeben, welche Felder indiziert werden sollen. Legen Sie die Werte wie folgt fest:



The screenshot shows the 'Full-Text Indexing Wizard' dialog box, specifically the 'Select Table Columns' step. The title bar reads 'Full-Text Indexing Wizard'. The main heading is 'Select Table Columns' with the instruction: 'Select the character-based or image-based columns you want to be eligible for full-text queries.' Below this is a table with four columns: 'Available Columns', 'Language for W...', 'Type Column', and 'Statistical Seman...'. The 'file_stream' row is selected with a checked checkbox, and its 'Language for W...' is set to 'German'. The 'file_type' row has an unchecked checkbox and 'file_type' in the 'Type Column' column. The 'name' row has an unchecked checkbox. At the bottom, there are buttons for 'Help', '< Back', 'Next >', 'Finish >>', and 'Cancel'.

Available Columns	Language for W...	Type Column	Statistical Seman...
<input checked="" type="checkbox"/> file_stream	German	file_type	<input type="checkbox"/>
<input type="checkbox"/> file_type			<input type="checkbox"/>
<input type="checkbox"/> name			<input type="checkbox"/>

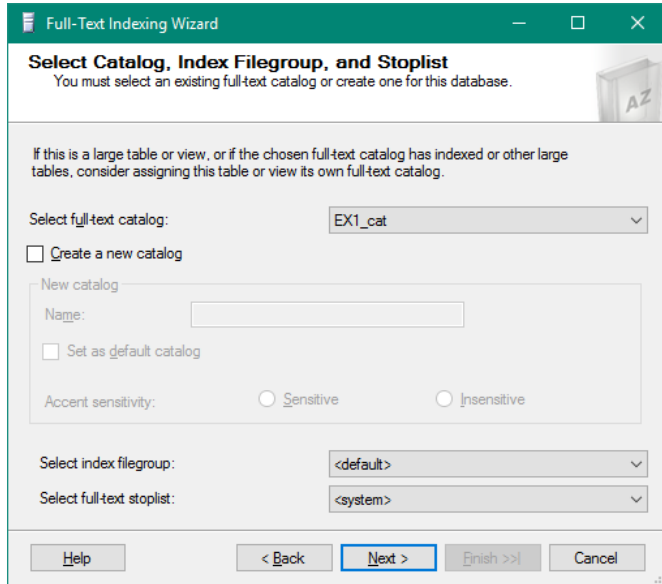
Hinweis: Die Sprachen sind zumindest in der englischen Version des SSMS nicht richtig sortiert. Sie finden German ziemlich weit unten, wenn Sie die Hoffnung schon fast aufgegeben haben.



The screenshot shows the 'Full-Text Indexing Wizard' dialog box, specifically the 'Select Change Tracking' step. The title bar reads 'Full-Text Indexing Wizard'. The main heading is 'Select Change Tracking' with the instruction: 'Choose how updates to tables and views are tracked.' Below this is a section titled 'Track changes on this table/view as they occur:' with three radio button options: 'Automatically' (selected), 'Manually', and 'Do not track changes'. There is also a checked checkbox for 'Start full population when index is created'. At the bottom, there is an information box with a blue 'i' icon and the text: 'When you define automatic or manual change tracking, a full population of the index occurs. To avoid a population at the end of this wizard, select the Do Not Track Changes option, and clear the Start Full Population When Index is Created checkbox.' At the bottom, there are buttons for 'Help', '< Back', 'Next >', 'Finish >>', and 'Cancel'.

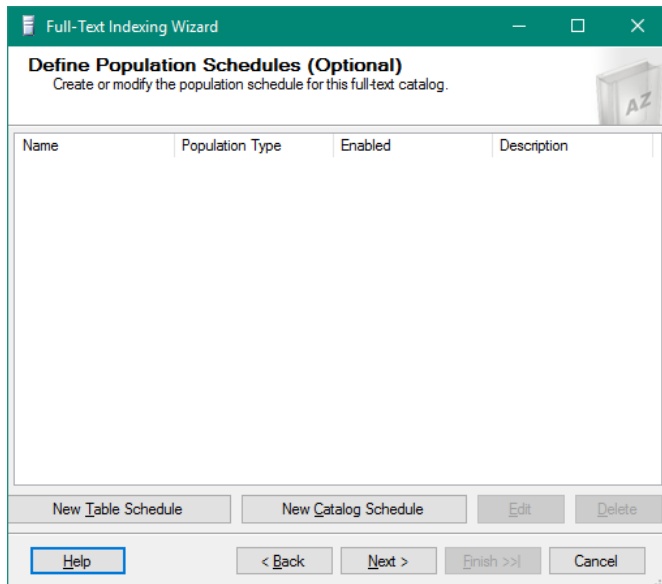


Den Katalog hatten wir ja bereits in einem früheren Schritt angelegt. Hier wäre jetzt nochmals die Gelegenheit dazu gewesen:



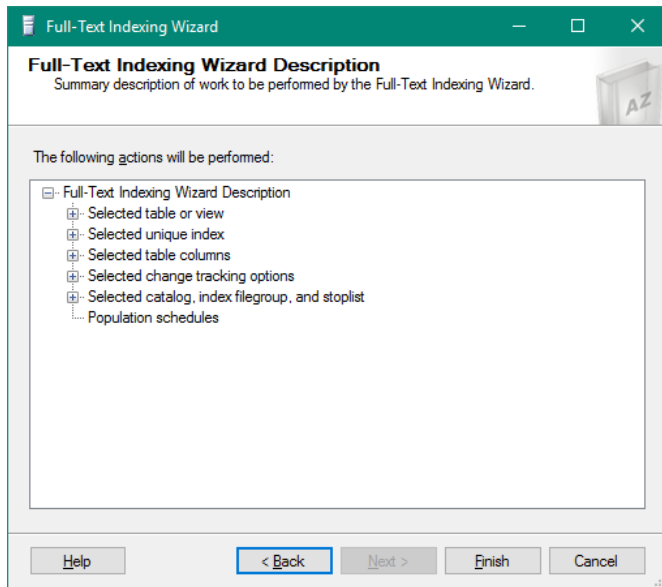
The screenshot shows the 'Full-Text Indexing Wizard' dialog box. The title bar reads 'Full-Text Indexing Wizard'. The main heading is 'Select Catalog, Index Filegroup, and Stoplist'. Below the heading, it says 'You must select an existing full-text catalog or create one for this database.' There is a small 'AZ' icon in the top right corner. The dialog contains several fields and options: 'Select full-text catalog:' with a dropdown menu showing 'EX1_cat'; a checkbox for 'Create a new catalog' which is unchecked; a section for 'New catalog' with a 'Name:' text box, a 'Set as default catalog' checkbox (unchecked), and 'Accent sensitivity:' options for 'Sensitive' and 'Insensitive' (both unselected); 'Select index filegroup:' with a dropdown menu showing '<default>'; and 'Select full-text stoplist:' with a dropdown menu showing '<system>'. At the bottom, there are buttons for 'Help', '< Back', 'Next >', 'Finish >>', and 'Cancel'. The 'Next >' button is highlighted with a blue border.

Mit Next bestätigen.

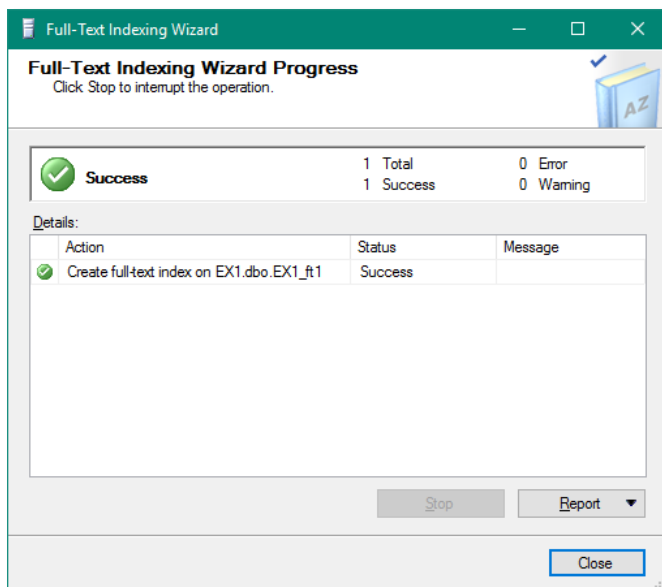


The screenshot shows the 'Full-Text Indexing Wizard' dialog box. The title bar reads 'Full-Text Indexing Wizard'. The main heading is 'Define Population Schedules (Optional)'. Below the heading, it says 'Create or modify the population schedule for this full-text catalog.' There is a small 'AZ' icon in the top right corner. The dialog contains a table with the following columns: 'Name', 'Population Type', 'Enabled', and 'Description'. The table is currently empty. Below the table, there are buttons for 'New Table Schedule', 'New Catalog Schedule', 'Edit', and 'Delete'. At the bottom, there are buttons for 'Help', '< Back', 'Next >', 'Finish >>', and 'Cancel'. The 'Help' button is highlighted with a blue border.

Mit Next bestätigen.



Mit Finish bestätigen.



Der Full-Text Index ist jetzt erfolgreich angelegt.



Die Volltextsuche

Nachdem die Vorarbeiten nun erledigt sind, können wir daran gehen, erste Schritte mit der Volltextsuche zu machen.

Die Erweiterung der SQL Sprache um Elemente für die Volltextsuche eröffnet umfassende Möglichkeiten, die den Rahmen dieser Einführung sprengen würden. Wir beschränken uns daher auf ein knappes Beispiel, das aber bereits den Umfang des Machbaren verdeutlicht.

Wir haben im Schritt 10 der Anleitung einige Dokumente in die Share kopiert. Zwei dieser Dokumente enthalten den Begriff „Wartung“ im Text.

Führen Sie folgendes Skript aus:

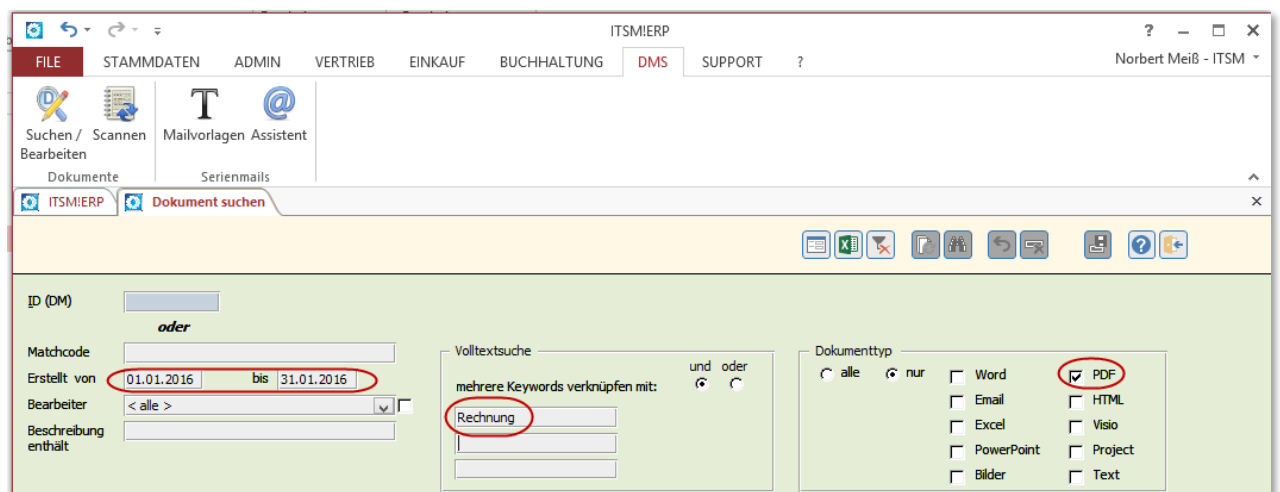
```
SELECT stream_id, name, file_type
FROM EX1_ft1
WHERE CONTAINS(file_stream, 'Wartung')
```

Als Ergebnis werden die beiden Dokumente angezeigt, die an beliebiger Stelle im Text den Begriff „Wartung“ enthalten:

	stream_id	name	file_type
1	10AB77A7-64D1-E511-9BD3-64006A62DCDB	KNX Reni.docx	docx
2	897FB2B2-65D1-E511-9BD3-64006A62DCDB	00118857.msg	msg

Dies ist zunächst nur ein sehr rudimentäres Ergebnis, mit dem ein Sachbearbeiter noch nicht allzu viel anfangen kann. Aber das ist ja auch nur der Anfang.

Natürlich können Sie die File Table Tabelle auch über Join mit anderen Tabellen verknüpfen und damit sinnvolle Darstellungen für den Benutzer schaffen. Hier ein Beispiel aus dem Modul Dokumentenmanagement von ITSM!ERP:



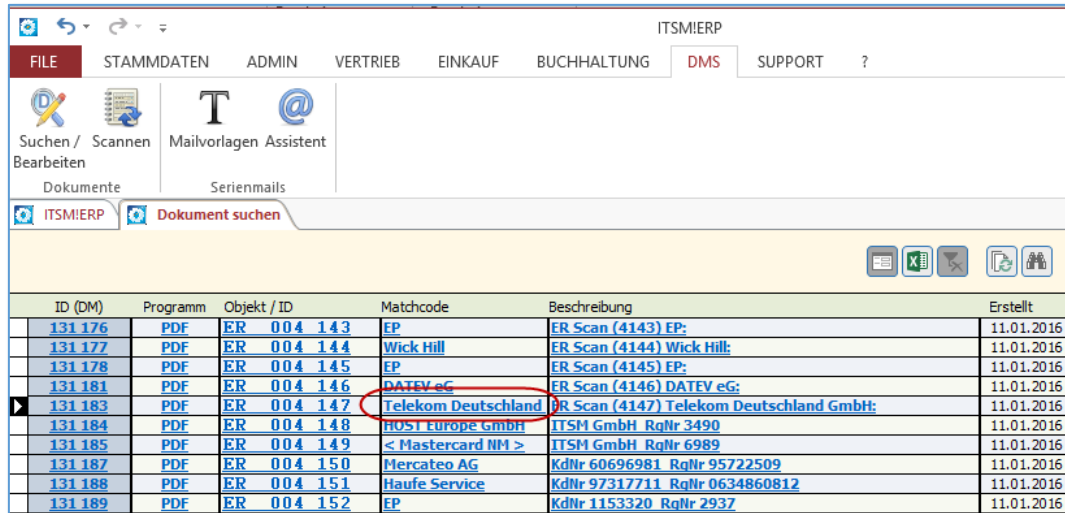
The screenshot shows the 'Dokument suchen' (Search Document) interface in the ITSM!ERP DMS module. The search criteria are as follows:

- ID (DM): [Empty]
- Matchcode: [Empty]
- Erstellt von: 01.01.2016 bis 31.01.2016
- Bearbeiter: < alle >
- Beschreibung enthält: [Empty]
- Volltextsuche: mehrere Keywords verknüpfen mit: Rechnung
- Dokumenttyp: alle, nur, Word, PDF (checked), Email, HTML, Excel, Visio, PowerPoint, Project, Bilder, Text

Der Suchauftrag lautet im Klartext: „Gib mir eine Liste aller PDFs aus dem Zeitraum Januar 2016, in denen das Wort ‚Rechnung‘ vorkommt“.



Die Ergebnisliste umfasst 77 Dokumente – das sind die bei ITSM eingegangenen und im DMS abgelegten Rechnungen, unter anderem eine gescannte Rechnung der Telekom:

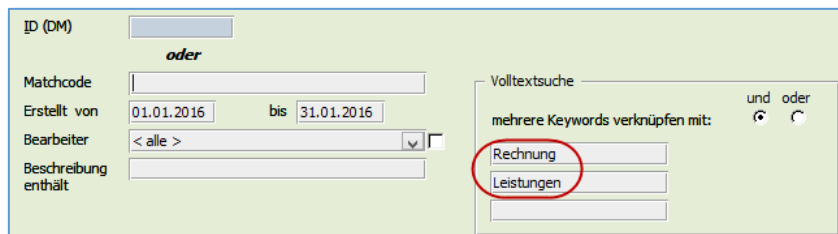


ID (DM)	Programm	Objekt / ID	Matchcode	Beschreibung	Erstellt
131 176	PDF	ER_004_143	EP	ER Scan (4143) EP:	11.01.2016
131 177	PDF	ER_004_144	Wick Hill	ER Scan (4144) Wick Hill:	11.01.2016
131 178	PDF	ER_004_145	EP	ER Scan (4145) EP:	11.01.2016
131 181	PDF	ER_004_146	DATEV eG	ER Scan (4146) DATEV eG:	11.01.2016
131 183	PDF	ER_004_147	Telekom Deutschland	ER Scan (4147) Telekom Deutschland GmbH:	11.01.2016
131 184	PDF	ER_004_148	HOST Europe GmbH	ITSM GmbH RqNr 3490	11.01.2016
131 185	PDF	ER_004_149	< Mastercard NM >	ITSM GmbH RqNr 6989	11.01.2016
131 187	PDF	ER_004_150	Mercateo AG	KdNr 60696981 RqNr 95722509	11.01.2016
131 188	PDF	ER_004_151	Haufe Service	KdNr 97317711 RqNr 0634860812	11.01.2016
131 189	PDF	ER_004_152	EP	KdNr 1153320 RqNr 2937	11.01.2016

Durch Klicken auf das Wort PDF in der Spalte Programm wird das PDF geöffnet. Die folgende Abbildung enthält den relevanten Abschnitt des Scans:



Der Ausschnitt des Scans enthält auch den Begriff „Leistungen“. Wenn wir diesen Suchbegriff mit eingeben, ...



ID (DM)

oder

Matchcode

Erstellt von 01.01.2016 bis 31.01.2016

Bearbeiter < alle >

Beschreibung enthält

Volltextsuche

mehrere Keywords verknüpfen mit: und oder

... wird die Ergebnisliste mit nur noch acht Einträgen kürzer. Sie enthält folgerichtig alle Rechnungen der Telekom, weil in allen das Wort „Leistungen“ vorkommt:



ID (DM)	Programm	Objekt / ID	Matchcode	Beschreibung	Erstellt
131 921	PDF	ER 004 175	Telekom Deutschland	ER Scan (4175) Telekom Deutschland GmbH:	20.01.2016
131 920	PDF	ER 004 174	Telekom Deutschland	ER Scan (4174) Telekom Deutschland GmbH:	20.01.2016
131 917	PDF	ER 004 173	Ingenieurbüro Stenzel:	ER Scan (4173) Ingenieurbüro Stenzel:	20.01.2016
131 900	PDF	ER 004 168	EP	ER Scan (4168) EP:	20.01.2016
131 890	PDF	ER 004 165	Telekom Deutschland	ER Scan (4165) Telekom Deutschland GmbH:	20.01.2016
131 184	PDF	ER 004 148	HOST Europe GmbH	ITSM GmbH Rqtr 3490	11.01.2016
131 183	PDF	ER 004 147	Telekom Deutschland	ER Scan (4147) Telekom Deutschland GmbH:	11.01.2016
130 948	PDF	ER 004 129	Haus- und Handwerke	ER Scan (4129) Haus- und Handwerkerservice:	06.01.2016

Wie Sie an diesem kurzen Beispiel erkennen können, lässt sich mit Hilfe der Volltextsuche, für die wir die File Tables im SQL Server benutzen, eine benutzerfreundliche Bedienoberfläche gestalten, die den Anwender mit wenigen Klicks zum gewünschten Dokument führt.

Wenn Sie mehr über die Möglichkeiten erfahren möchten, buchen Sie bei uns einen Workshop zu diesem Thema. Wir zeigen Ihnen die Volltextsuche in der Praxis und loten mit Ihnen gemeinsam aus, wie Sie dieses mächtige Feature in Ihrem Unternehmen sinnvoll einsetzen können.



Über ITSM

Die [ITSM Gesellschaft für Informationstechnologie und Services Meiß mbH](#) ist ein 1998 gegründetes IT-Systemhaus in Langenfeld/Rheinland, das mittelständische Unternehmen in allen Fragen der Informationsverarbeitung und Telekommunikation berät und unterstützt. Die Gesellschaft bedient schwerpunktmäßig die Themen „Server Based Computing“ und „Local Cloud Hosting“, die erhebliches Einsparungspotenzial bei Investitionen und laufendem Betrieb aufweisen. ITSM plant und realisiert als Microsoft Partner individuell ausgelegte Netzwerkinfrastrukturen, die auch in deutschen Rechenzentren gehostete Server und Speicher umfassen.

In 2015 wurde die Softwareentwicklung in die Schwestergesellschaft ITSM OHG ausgegliedert. Hier erfolgt die kontinuierliche Weiterentwicklung des Kernprodukts ITSM!ERP. Neben dieser individuell anpassbaren ERP/CRM-Lösung auf Basis der Produkte Microsoft SQL Server und Microsoft Office bietet die ITSM OHG auch Auftragsprogrammierung wie zum Beispiel Schnittstellen oder Webservices an.

Beide Unternehmen arbeiten an einem gemeinsamen Standort in Langenfeld/Rheinland – also in zentraler Lage zwischen Düsseldorf und Köln – eng zusammen und stehen unter gemeinsamer Leitung. Seit Jahren bildet ITSM selber Fachinformatiker mit dem Ziel der späteren Übernahme als feste Mitarbeiter aus. Anfang 2016 beschäftigt ITSM 15 Mitarbeiter.

75124_00117086-0002.docx