DB2 Universal Database



Text Extender: Verwaltung und Programmierung

Version 6

SC12-2709-00

DB2 Universal Database



Text Extender: Verwaltung und Programmierung

Version 6

SC12-2709-00

Anmerkung

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts, sollten die allgemeinen Informationen unter „Bemerkungen“ auf Seite 275 gelesen werden.

Erste Ausgabe (Juni 1999)

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
DB2 Universal Database Text Extender: Administration and Programming,
IBM Form SC26-9646-00,

herausgegeben von International Business Machines Corporation, USA

(C) Copyright International Business Machines Corporation 1999

(C) Copyright IBM Deutschland Informationssysteme GmbH 1999

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderungen des Textes bleiben vorbehalten.

Herausgegeben von:
SW NLS Center
Kst. 2877
Juni 1999

Inhaltsverzeichnis

Zusammenfassung der Änderungen xiii

Zu diesem Buch xv
Zielgruppe xv
Aufbau dieses Buchs xv
Beschreibung der Syntaxdiagramme xvii
Zugehörige Informationen xviii
Kommentare zum vorliegenden Buch xx

Teil 1. Einführung 1

Kapitel 1. Ein Überblick über die DB2 Extender 3
Text Extender 3
Andere DB2 Extender 4
Image Extender 5
Video Extender 5
Audio Extender 6
Der Text Extender in der DB2 Client/Server-Umgebung 7
Der Text Extender in der partitionierten Datenbankumgebung 9

Kapitel 2. Einen Textindex planen 11
Warum Textdokumente indexiert sein müssen 11
Indextypen 13
Linguistischer Index 13
Präziser Index 14
Dualer Index 15
Ngram-Index 15
Den Textindextyp ändern 16
Einen oder mehrere Textindizes für eine Tabelle erstellen 16
Die Größe eines Indexes berechnen 17

Kapitel 3. Verarbeitung auf linguistischer Basis 19
Verarbeitung auf linguistischer Basis beim Indexieren 19
Basistextanalyse 21
Begriffe auf ihre Grundform reduzieren 22
Stoppwortfiltern 23
Dekomposition (Trennen von Komposita) 23
Verarbeitung auf linguistischer Basis für die Abfrage 24
Synonyme 25
Thesaurus-Erweiterung 26
Lauterweiterung 26
Platzhalterzeichen für Zeichen und Wörter 26
Verarbeitung auf linguistischer Basis für die Anzeige 27
Stufe 1: Normalisierung und Begriffserweiterung 27

Stufe 2: Erweiterte Übereinstimmung 28
Linguistische Funktionen für die unterstützten Sprachen 30
Wörterverzeichnisse, Stoppwortlisten und Abkürzungslisten 31
Thesaurus-Konzepte 32
Begriffe 34
Relationen 34
Ngram-Thesaurus-Relation 36
Einen Thesaurus erstellen 37
! Einen Ngram-Thesaurus erstellen 41

Kapitel 4. Verwaltung 43
! Der Beispiel-Wizard für die Indexierung und die Suche (nur NT) 43
Ein Text Extender-Exemplar erstellen 43
Einen Text Extender-Server verwalten 44
Den Server starten 44
Den Status des Servers anzeigen 44
Den Server stoppen 44
Überblick über die Verwaltungs-Tasks auf dem Client 45
Verwaltungsüberblick 45
Einführung 47
Die Verwaltung starten 48
Den Text Extender-Befehlszeilenprozessor starten (wahlfrei) 48
Eine Verbindung zu einer Datenbank herstellen 49
Textdokumente für die Suche vorbereiten 50
Die Textkonfiguration ändern 50
Die Stoppwort- und Abkürzungsdateien ändern 52
! Die Dokumentmodelldatei ändern 53
Eine Beispieltabelle erstellen 54
Eine Datenbank aktivieren 56
Eine Texttabelle aktivieren 57
Eine Textspalte aktivieren 60
Externe Textdateien aktivieren 72
Die Verwaltungssitzung beenden 74
Den Textvorbereitungsprozeß umkehren 74
Eine Textspalte inaktivieren 74
Textdateien inaktivieren 75
Eine Texttabelle inaktivieren 76
Eine Datenbank inaktivieren 76
Textindizes verwalten 77
Einen Index aktualisieren 77
Einen Index für externe Dateien aktualisieren 78
Die Einstellungen eines Indexes ändern 79
Den Indexstatus zurücksetzen 79
Indexereignisse löschen 80
! Einen Index reorganisieren 81
Nützliche Informationen abrufen 82
Informationen zum aktivierten Status anzeigen 82
Die Einstellungen der Umgebungsvariablen anzeigen 83

Die Textkonfigurationseinstellungen anzeigen 84
Den Status eines Indexes anzeigen 84
Fehlerereignisse anzeigen 86
Die Indexeinstellungen anzeigen 87
Die Texteinstellungen für eine Spalte anzeigen 88
Mit der Text Extender-Katalogsicht arbeiten 89
Die Trace-Funktion ausführen 91
Indizes und aktivierte Datenbanken sichern und zurückschreiben 93

Kapitel 5. Mit Hilfe von Text Extender-UDFs suchen 95
Die Beispiel-UDFs 95
Die Beispieltabelle DB2TX.SAMPLE 96
Kennungen für externe Dateien 99
Den aktuellen Funktionspfad setzen 99
Nach Text suchen 100
Eine Abfrage durchführen 101
Die Anzahl der gefundenen Übereinstimmungen suchen und zurückgeben . . 102
Die Rangordnung eines gefundenen Textdokuments suchen und zurückgeben 102
Suchargumente angeben 102
Mehrere Begriffe suchen 102
Mit den Booleschen Operatoren AND und OR suchen 103
Nach Variationen eines Begriffs suchen 104
Nach Teilen eines Begriffs suchen (mit Platzhalterzeichen) 104
Nach Begriffen suchen, die bereits ein Platzhalterzeichen enthalten 105
Nach Begriffen in beliebiger Reihenfolge suchen 106
Nach Begriffen im selben Satz oder Absatz suchen 106
! Nach Begriffen in Abschnitten von strukturierten Dokumenten suchen 106
Nach Synonymen von Begriffen suchen 106
Eine linguistische Suche durchführen 107
Mit dem Booleschen Operator NOT suchen 108
Suche nach grober Übereinstimmung 109
Wortgrenzen respektieren 109
Nach ähnlich klingenden Wörtern suchen 109
Thesaurus-Suche 110
Freitext- und Hybridsuche 111
Eine Feinanpassung für eine vorherige Suche durchführen 111
Informationen in Kennungen setzen und extrahieren 113
Textdaten beim Einfügen von neuem Text setzen 113
Informationen aus Kennungen extrahieren 114
Informationen in Kennungen ändern 114
Die Suchleistung verbessern 115

Kapitel 6. Die API-Funktionen zum Suchen und Anzeigen verwenden 117
Ihre Anwendung einrichten 117
Eine UNIX-Anwendung verbinden 117
Eine OS/2- oder Windows-Anwendung verbinden 117
Überblick über die API-Funktionen 118
Nach Text suchen 118

Text anzeigen 119
Nach Text suchen 120
Eine Suchergebnistabelle abrufen (DesGetSearchResultTable) 120
Text anzeigen 121
Anzeigeinformationen abrufen (DesGetBrowseInfo) 121
Eine Anzeigesitzung starten (DesStartBrowseSession) 121
Ein Dokument öffnen (DesOpenDocument) 122
Übereinstimmungen abrufen (DesGetMatches) 122
Ein Dokument schließen (DesCloseDocument) 123
Eine Anzeigesitzung beenden (DesEndBrowseSession) 123
Die Anzeigeinformationen freigeben (DesFreeBrowseInfo) 123

Teil 2. Referenz 125

Kapitel 7. Verwaltungsbefehle für den Client 127
Hilfe für den Befehlszeilenprozessor 128
CHANGE INDEX SETTINGS 129
CHANGE TEXT CONFIGURATION 130
CONNECT 133
DELETE INDEX EVENTS 134
DISABLE DATABASE 135
DISABLE TEXT COLUMN 136
DISABLE TEXT FILES 137
DISABLE TEXT TABLE 138
ENABLE DATABASE 139
ENABLE TEXT COLUMN 139
ENABLE TEXT FILES 147
ENABLE TEXT TABLE 150
GET ENVIRONMENT 154
GET INDEX SETTINGS 154
GET INDEX STATUS 156
GET STATUS 157
GET TEXT CONFIGURATION 157
GET TEXT INFO 158
QUIT 158
! REORGANIZE INDEX 159
RESET INDEX STATUS 160
UPDATE INDEX 160

Kapitel 8. Verwaltungsbefehle für den Server 163
TXICRT 163
TXIDROP 165
TXILIST (nur UNIX) 165
TXNADD 166
TXNCHECK 167
TXNDROP 167
TXSAMPLE 168
TXSTART 168
TXSTATUS 169
TXSTOP 169
! TXTHESC 169
! TXTHESN 171
TXTRACE 172
TXVERIFY 178

Kapitel 9. UDTs und UDFs 179
Eine Zusammenfassung der Text Extender-UDTs 179
Eine Zusammenfassung der Text Extender-UDFs 181
CCSID 182
CONTAINS 182
FILE 183
FORMAT 184
HANDLE 185
HANDLE_LIST 186
INIT_TEXT_HANDLE 187
LANGUAGE 188
NO_OF_DOCUMENTS 189
NO_OF_MATCHES 189
RANK 190
REFINE 190
SEARCH_RESULT 191

Kapitel 10. Syntax von Suchargumenten 193
Suchargument 194

Kapitel 11. API-Funktionen zum Suchen und Anzeigen 205
DesCloseDocument 206
DesEndBrowseSession 207
DesFreeBrowseInfo 208
DesGetBrowseInfo 208
DesGetMatches 211
DesGetSearchResultTable 216
DesOpenDocument 220
DesStartBrowseSession 222

Kapitel 12. Rückkehrcodes 225

Kapitel 13. Nachrichten 235
SQLSTATES, die von UDFs zurückgegeben werden 235
Nachrichten vom Text Extender 238

Kapitel 14. Konfiguration 253
Umgebungsvariablen 253
Textkonfigurationseinstellungen 254
Textmerkmale 254
Indexmerkmale 255
Verarbeitungsmerkmale 255
Informationen zu Textdokumenten 256
Formate 256
Sprachen 258
! CCSIDs 259
Die Frequenz von Indexaktualisierungen setzen 262

Kapitel 15. Beispiel-API-Programm 265

Kapitel 16. Ursachencodes für Fehlerereignisse 267

Bemerkungen 275
Informationen zur Programmierschnittstelle 276
Marken 276

Glossar 277

Index 283

Abbildungsverzeichnis

1. Integration des Text Extender in der DB2 Client/Server-Umgebung 7
2. Eine Text Extender-Konfiguration 8
3. Integration des Text Extender in die DB2-Umgebung für partitionierte Datenbanken 9
4. Nur wesentliche Begriffe werden indexiert 12
5. Ein als Netz angezeigter Thesaurus 33
6. Die Definition eines einfachen Thesaurus 39
7. Eine Datenbank aktivieren 56
8. Einen gemeinsamen Index für alle Textspalten in einer Tabelle erstellen . 59
9. Einen separaten Index für jede Textspalte erstellen 60
10. Struktur der Tabelle DB2TX.SAMPLE - vor dem Aktivieren 61
11. Struktur der Tabelle DB2TX.SAMPLE - nach dem Aktivieren 61
12. Die Struktur der Tabelle DB2TX.SAMPLE 97
13. Die Tabelle DB2TX.SAMPLE nach dem Aktivieren 98
14. Die Kennung für eine eingefügte Zeile wird von einem Auslöser erstellt . . 98
15. Reihenfolge für die Verwendung der API-Funktionen 119
16. Struktur der Ergebnistabelle 218

Tabellen

1. Begriffsextraktion für einen linguistischen Index 20
2. Begriffsextraktion für einen präzisen Index 20
3. Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden 30
4. Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden 31
5. Text Extender-Katalogsicht 89
6. Ein Auszug aus der Beispieltabelle DB2TX.SAMPLE 96
7. Linguistische Optionen 198
8. Suchbegriffsoptionen für Ngram-Indizes 199
9. Argumente für DesCloseDocument 206
10. Argumente für DesEndBrowseSession 207
11. Argumente für DesFreeBrowseInfo 208
12. Argumente für DesGetBrowseInfo 209
13. Argumente für DesGetMatches 212
14. Argumente für DesGetSearchResultTable 216
15. Argumente für DesOpenDocument 220
16. Argumente für DesStartBrowseSession 222

Zusammenfassung der Änderungen

! **Die Unterstützung strukturierter Dokumente** wurde hinzugefügt. Der Text Extender
! kann die Struktur von codierten Dokumenten auf der Basis eines Dokumentmodells
! erkennen, das Sie liefern. Die unterstützten codierten Dokumente sind XML-, HTML-
! und codierte ASCII-Dokumente. Der Vorteil dieser Funktion liegt darin, daß eine Suche
! auf einen bestimmten Abschnitt oder eine Liste von Abschnitten beschränkt werden
! kann.

! **Die linguistische Unterstützung für Arabisch, Hebräisch und Russisch** wurde
! hinzugefügt.

! **Die Anzahl der unterstützten CCSIDs wurde erweitert.** Dokumente in Unicode
! können jetzt indexiert und durchsucht werden.

! **Für Windows NT-Benutzer wird ein Beispiel für eine Verwaltungs- und**
! **Suchschnittstelle geboten.** Die Schnittstelle wird in Form eines Wizard geliefert, der
! Sie durch die Tasks zur Verwaltung und Suche leitet. Der Java-Quellencode wird zur
! Verfügung gestellt.

! **Textspalten können mehrfach aktiviert werden**. Hierdurch können Sie Textspalten
! für mehr als einen Indextyp aktivieren oder verschiedene Indexoptionen verwenden.

! **Die Angabe eines Tabellenbereichs** für Protokolltabellen ist möglich.

! Ein neuer Verwaltungsbefehl wurde hinzugefügt für die **Reorganisation eines Indexes**.
! Diese Funktion wird zusätzlich zur aktuellen automatischen Reorganisation angeboten.
! Nachdem ein Index häufig aktualisiert wurde, werden durch die Indexreorganisation von
! Indizes die Suchen viel effizienter.

! Der Befehl **GET INDEX STATUS** zeigt jetzt den Status des Primär- und des Sekundär-
! indexes an, so daß Sie bestimmen können, ob die manuelle Reorganisation eines
! Indexes erforderlich ist.

! Sie können jetzt **einen Ngram-Thesaurus kompilieren**, indem Sie den neuen Befehl
! TXTHESN verwenden. Dieser Befehl ergänzt den Befehl TXTHESC, mit dem ein
! Standard-Thesaurus kompiliert wird.

Zu diesem Buch

DB2 Text Extender ist ein Volltextabfrageprogramm für das Suchen in Textdateien, die in einer DB2 Universal Database-Datenbank gespeichert sind, aber auch für das Suchen in Textdateien, die an anderen Orten gespeichert sind. Das Programm liefert Erweiterungen zur Structured Query Language (SQL), die mit einer leistungsfähigen und intelligenten Suchmaschine arbeiten, so daß Sie schnell und effizient in unstrukturiertem Text nach Informationen suchen können.

In diesem Buch wird beschrieben, wie der Text Extender zur Vorbereitung und Verwaltung einer DB2-Datenbank für Textdaten verwendet wird. Außerdem wird beschrieben, wie Sie die benutzerdefinierten Funktionen (UDFs) und die Anwendungsprogrammierschnittstellen (APIs) verwenden können, die vom Text Extender für den Zugriff und die Bearbeitung dieser Daten zur Verfügung gestellt werden. Verweise auf "DB2" in diesem Buch beziehen sich auf DB2 UDB Version 6 oder höher.

Zielgruppe

Dieses Buch ist für DB2-Datenbankadministratoren konzipiert, die mit den Konzepten, Tools und Methoden der DB2-Verwaltung vertraut sind.

Dieses Buch ist außerdem für DB2-Anwendungsprogrammierer konzipiert, die mit SQL und einer oder mehreren Programmiersprachen vertraut sind, die für DB2-Anwendungsprogramme verwendet werden können.

Das Buch ist für Personen gedacht, die mit dem Text Extender arbeiten werden. Personen, die mit den DB2 Image, Audio und Video Extender arbeiten, sollten das Buch *DB2 Image, Audio und Video Extender Verwaltung und Programmierung* lesen.

Aufbau dieses Buchs

Dieses Buch enthält die folgenden Kapitel:

Teil 1: Einführung

Kapitel 1, Ein Überblick über die DB2 Extender
Stellt den Text Extender vor und beschreibt andere Produkte in der DB2 Extender-Produktfamilie.

Kapitel 2, Einen Textindex planen
Hilft Ihnen bei der Entscheidung, welcher Typ von Textindex zu Ihren Anforderungen paßt.

Text Extender-Administratoren benötigen diese Informationen beim Einrichten des Standardindextyps in den Textkonfigurationsdaten.

Text Extender-Benutzer benötigen diese Informationen beim Aktivieren einer Texttabelle bzw. einzelner Textspalten für die Suche.

Aufbau dieses Buchs

Kapitel 3, Verarbeitung auf linguistischer Basis
Erläutert, wie der Text Extender Dokumentbegriffe beim Indexieren, Suchbegriffe beim Suchen und übereinstimmende Begriffe beim Hervorheben in einem gefundenen Dokument analysiert.

Text Extender-Benutzer benötigen diese Informationen beim Erstellen effektiverer Suchbegriffe.

Kapitel 4, Verwaltung
Beschreibt den Text Extender-Befehlszeilenprozessor.

Text Extender-Administratoren benötigen diese Informationen beim Einrichten eines Text Extender-Servers.

Text Extender-Benutzer benötigen diese Informationen zum Vorbereiten von Tabellen und einzelnen Textspalten zur Verwendung durch den Text Extender.

Kapitel 5, Mit Hilfe von Text Extender-UDFs suchen
Beschreibt die benutzerdefinierten SQL-Funktionen (UDFs), die vom Text Extender geliefert werden.

Text Extender-Benutzer benötigen diese Informationen beim Einschließen von Unterabfragen zur Textsuche in SQL-Abfragen.

Kapitel 6, Die API-Funktionen zum Suchen und Anzeigen verwenden
Beschreibt die API-Funktionen (API = Anwendungsprogrammierschnittstelle), die vom Text Extender geliefert werden.

Text Extender-Benutzer benötigen diese Informationen beim Einschließen von Unterabfragen zur Textsuche in Anwendungsprogrammen.

Teil 2: Referenz

Kapitel 7, Verwaltungsbefehle für den Client Die Syntax und Beschreibung der Verwaltungsbefehle für den Client in alphabetischer Reihenfolge.

Kapitel 8, Verwaltungsbefehle für den Server Die Syntax und Beschreibung der Verwaltungsbefehle für den Server in alphabetischer Reihenfolge.

Kapitel 9, UDTs und UDFs Die Beschreibung der UDTs, die vom Text Extender geliefert werden, und die Syntax und Beschreibung der Text Extender-UDFs in alphabetischer Reihenfolge.

Kapitel 10, Syntax von Suchargumenten Die Syntax von Suchargumenten, die in SQL-Abfragen und in den API-Funktionen verwendet werden.

Kapitel 11, API-Funktionen zum Suchen und Anzeigen Die Syntax und Beschreibung der API-Funktionen in alphabetischer Reihenfolge.

Kapitel 12, Rückkehrcodes Eine Erläuterung der Rückkehrcodes von den API-Funktionen.

Kapitel 13, Nachrichten Eine Erläuterung der Nachrichten von den UDFs und vom Verwaltungsbefehlszeilenprozessor.

Kapitel 14, Konfiguration Eine Beschreibung der Umgebungsvariablen für den Text Extender und der Textkonfiguration, die die Standardkonfigurationsdaten enthält.

Kapitel 15, Beispiel-API-Programm Ein Beispielprogramm in C, das die API-Funktionen verwendet.

Kapitel 16, Ursachencodes für Fehlerereignisse Eine Beschreibung der Ursachencodes bei Fehlern, die während der Indexierung auftreten können.

Beschreibung der Syntaxdiagramme

Die Syntax, die in diesem Buch beschrieben wird, ist wie folgt strukturiert:

- Syntaxdiagramme werden von links nach rechts und von oben nach unten gelesen, wobei die Linie die Richtung angibt.

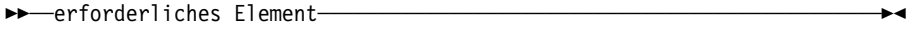
  Das Symbol ▸▸— kennzeichnet den Anfang einer Anweisung.

  Das Symbol —▸ gibt an, daß die Anweisungssyntax in der nächsten Zeile fortgesetzt wird.

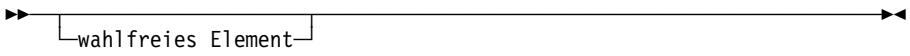
  Das Symbol ▸— gibt an, daß es sich um die Fortsetzung einer Anweisung aus der vorherigen Zeile handelt.

  Das Symbol —▸◂ kennzeichnet das Ende einer Anweisung.

- Erforderliche Elemente sind auf der horizontalen Linie (dem Hauptpfad) dargestellt.

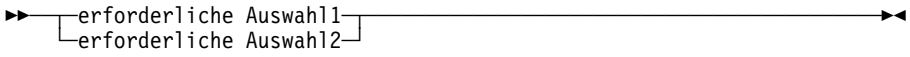
  

- Wahlfreie Elemente werden unterhalb des Hauptpfads dargestellt.

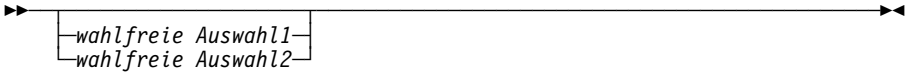
- Besteht die Möglichkeit, aus mehreren Elementen auszuwählen, sind diese untereinander angeordnet.

  Wenn Sie eines dieser Elemente auswählen *müssen*, ist eines der untereinander angeordneten Elemente auf dem Hauptpfad dargestellt.

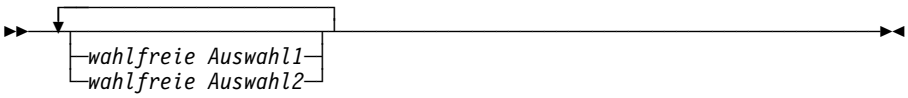
  

Zugehörige Informationen

Ist es nicht unbedingt erforderlich, daß Sie eines der Elemente auswählen, werden alle auswählbaren Elemente untereinander unterhalb des Hauptpfads dargestellt.



Ein Wiederholungspfeil oberhalb der aufgelisteten Elemente gibt an, daß Sie mehrere der aufgeführten Elemente auswählen können.



- Schlüsselwörter sind in Großbuchstaben dargestellt, ihre Schreibweise muß exakt eingehalten werden. Variablen sind in Kleinbuchstaben dargestellt (z.&rbl;B. `srcpath`). Mit den Variablen werden in der Syntax Namen oder Werte dargestellt, die vom Benutzer eingegeben werden.
- Sind Interpunktionszeichen, runde Klammern, arithmetische Operatoren oder andere, ähnliche Symbole dargestellt, müssen diese als Teil der Syntax eingegeben werden.

Zugehörige Informationen

DB2 Universal Database Version 6

DB2 Universal Database Einstieg, Version 6, IBM Form GC12-2696 (OS/2), GC12-2697 (Windows NT), GC12-2698 (UNIX). In diesen Büchern wird die Planung, Installation, Konfiguration und Migration der DB2 Universal Database auf den entsprechenden Plattformen beschrieben.

DB2 Universal Database Enterprise-Extended Edition Quick Beginnings, Version 6, IBM Form GC09-2832 (AIX), GC09-2833 (Windows NT). In diesen Büchern wird die Planung, Installation und Konfiguration der DB2 Universal Database Enterprise-Extended Edition auf den entsprechenden Plattformen beschrieben.

DB2 Universal Database Systemverwaltung: Konzept und Implementierung, Version 6, Band 1, IBM Form SC12-2702. In diesen Buch wird beschrieben, wie eine DB2-Datenbank konzipiert und implementiert wird.

DB2 Universal Database Systemverwaltung: Leistung, Version 6, Band 2, IBM Form SC12-2703. In diesem Buch wird beschrieben, wie die Leistung einer DB2-Datenbank optimiert wird.

DB2 Universal Database Application Development Guide, Version 6, IBM Form SC09-2845. In diesem Buch wird beschrieben, wie Anwendungen entwickelt werden, die unter Verwendung von eingebettetem SQL oder JDBC auf DB2-Datenbanken zugreifen.

Außerdem wird beschrieben, wie gespeicherte Prozeduren (Stored Procedures) und benutzerdefinierte Funktionen geschrieben werden, wie benutzerdefinierte Typen definiert werden und wie Auslöser verwendet werden.

DB2 Universal Database CLI Guide and Reference, Version 6, IBM Form SC09-2843. In diesem Buch wird beschrieben, wie Anwendungen entwickelt werden, die unter Verwendung der DB2 Call Level Interface, einer aufrufbaren SQL-Schnittstelle, die mit der Microsoft-ODBC-Spezifikation kompatibel ist, auf DB2-Datenbanken zugreifen.

DB2 Universal Database Command Reference, Version 6, IBM Form SC09-2844. In diesem Buch wird beschrieben, wie der DB2-Befehlszeilenprozessor verwendet wird, und werden Referenzinformationen zu DB2-Befehlen gegeben.

DB2 Universal Database Fehlernachrichten, Version 6, IBM Form GC12-2699. In diesem Buch werden die Nachrichten und Codes, die von DB2 verwendet werden, aufgelistet und beschrieben und werden die Aktionen beschrieben, mit denen ein Benutzer das System nach einem bestimmten Fehler oder Problem wiederherstellen kann.

IBM DB2 Universal Database Enterprise-Extended Edition, Version 6

IBM DB2 Universal Database Enterprise-Extended Edition Quick Beginnings, Version 6, IBM Form GC09-2832 (UNIX), GC09-2833 (Windows NT). In diesen Büchern wird die Planung, Installation und Konfiguration der DB2 Universal Database Enterprise-Extended Edition auf den entsprechenden Plattformen beschrieben.

DB2 Universal Database Image, Audio und Video Extender, Version 6

DB2 Universal Database Image, Audio und Video Extender Verwaltung und Programmierung, Version 6, IBM Form SC12-2708. In diesem Buch wird beschrieben, wie eine DB2-Datenbank für Abbild-, Audio- und Videodaten verwaltet wird. Außerdem wird die Verwendung der Anwendungsprogrammierschnittstellen beschrieben, die mit den DB2 Image, Audio und Video Extendern geliefert werden, um auf diese Daten zuzugreifen und sie zu verwalten.

DB2 Universal Database for OS/390 Image, Audio, and Video Extenders, Version 6

DB2 Universal Database for OS/390 Version 6 Image, Audio, and Video Extenders Administration and Programming, IBM Form SC26-9650. In diesem Buch wird beschrieben, wie ein DB2 für OS/390-Datenbank-Server für Abbild-, Audio- und Videodaten verwaltet wird. Außerdem wird die Verwendung von benutzerdefinierten Funktionen und Anwendungsprogrammierschnittstellen beschrieben, die von den DB2 for OS/390 Image, Audio und Video Extendern geliefert werden, um auf Abbild-, Audio- und Videodaten zuzugreifen und diese Daten zu bearbeiten.

DB2 Universal Database for OS/390 Text Extender, Version 6

DB2 Universal Database for OS/390 Version 6 Text Extender Administration and Programming, IBM Form SC26-9651. In diesem Buch wird beschrieben, wie ein DB2 for OS/390-Datenbank-Server für Textdaten verwaltet wird. Außerdem wird die Verwendung von benutzerdefinierten Funktionen und Anwendungsprogrammierschnittstellen beschrieben, die vom DB2 for OS/390 Text Extender geliefert werden, um auf Textdaten zuzugreifen und diese Daten zu bearbeiten.

Die DB2 Extender-Web-Site. Diese Web-Site enthält Informationen zu den DB2 Extendern. Die URL-Adresse der DB2 Extender-Home Page lautet: http://www.software.ibm.com/data/db2/extenders.

Kommentare zum vorliegenden Buch

Ihre Rückmeldung ist für uns wichtig, um Ihnen genaueste Informationen in hoher Qualität zu liefern. Wenn Sie Kommentare und Anregungen zu diesem Buch haben:

- Besuchen Sie unsere Home Page unter:

  `http://www.software.ibm.com/data/db2/extenders`

  Dort finden Sie eine Antwortseite, auf der Sie Kommentare und Anregungen eingeben und an uns senden können.

- Senden Sie Ihre Kommentare als E-mail an swsdid@de.ibm.com oder an die IBMMAIL-Adresse DEIBM3P3@IBMMAIL. Geben Sie bitte den Namen des Buchs, die Teilenummer des Buchs, die Version des Produkts und, falls möglich, die Textstelle an, auf die sich der Kommentar bezieht (z.&rbl;B. Seitennummer oder Tabellennummer).

- Schicken Sie Kommentare an:

  IBM Corporation
  Department W92/H3
  P.O. Box 49023
  San Jose, CA 95161-9023
  U.S.A.

  Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne daß eine Verpflichtung gegenüber dem Einsender entsteht.

- Füllen Sie eines der Formulare am Ende dieses Buchs aus und schicken Sie es per Post oder per Fax zurück oder geben Sie es einem IBM Ansprechpartner. Die Postadresse befindet sich auf der Rückseite des Antwortformulars. Die Faxnummer ist +49-(0)7031-16-6901.

Teil 1. Einführung

Kapitel 1. Ein Überblick über die DB2 Extender

Der Text Extender ist ein Mitglied der DB2 Extender-Produktfamilie. Er ermöglicht es Programmierern, SQL-Abfragen für Textdokumente in ihre Anwendungen einzuschließen.

Die anderen Extender sind in der Lage, nach Abbild-, Video- und Sprachdaten zu suchen.

Text Extender

Der Text Extender integriert das Leistungsmerkmal der Volltextabfrage in SQL-Abfragen, indem Funktionen verwendet werden, die unter DB2 UDB zur Verfügung stehen. Mit diesen Funktionen können unstrukturierte Textdokumente von bis zu 2 Gigabyte in Datenbanken gespeichert werden.

Der Text Extender bietet DB2 UDB-Benutzern und Anwendungsprogrammierern eine schnelle, vielseitige und intelligente Methode zum Durchsuchen von Textdokumenten. Die Stärke des Text Extender liegt in seiner Fähigkeit, Tausende von großen Textdokumenten mit hoher Geschwindigkeit zu durchsuchen, wobei er nicht nur den direkten Suchbegriff findet, sondern auch Wortvariationen und Synonyme.

Sie sind nicht darauf beschränkt, nur in Textdokumenten zu suchen, die in DB2 UDB-Datenbanken gespeichert sind, sondern Sie können auch in Textdokumenten suchen, die in Dateien gespeichert sind.

Der Text Extender kann auf jede Art von Textdokument zugreifen, auch auf Textverarbeitungsdokumente in ihren ursprünglichen Basisformaten. Er bietet eine umfangreiche Palette von Abruffähigkeiten, einschließlich die Suche nach Wörtern, Ausdrücken, Platzhalterzeichen sowie die verknüpfte Suche unter Verwendung der Booleschen Logik.

Im Zentrum des Text Extender steht die IBM Hochleistungstechnologie für die linguistische Suche. Sie ermöglicht Ihren Anwendungen, auf unterschiedlichste Weise auf Textdokumente zuzugreifen und sie abzurufen. Ihre Anwendung kann folgendes:

- Suchen nach Dokumenten, die bestimmten Text, Synonyme eines Worts oder Ausdrucks oder Suchbegriffe in der Nähe, z.&rbl;B. im selben Satz oder Absatz, enthalten.
- Suchen mit Platzhalterzeichen, wobei das Platzhalterzeichen am Anfang, in der Mitte oder am Ende des Worts verwendet werden kann, bzw. als Platzhalterzeichen für ein Wort oder ein Zeichen.
- Suchen nach Dokumenten in verschiedenen Sprachen in verschiedenen Dokumentformaten.
- Ausführen einer Suche nach grober Übereinstimmung (“fuzzy”) nach Wörtern, die ähnlich geschrieben werden wie der Suchbegriff. Diese Art der Suche ist sinnvoll beim Suchen nach Wörtern, die möglicherweise Rechtschreibfehler enthalten.

Andere DB2 Extender

- Ausführen einer Freitextsuche, in der das Suchargument in natürlicher Sprache ausgedrückt wird.
- Suchen nach Namen von Personen, Orten und Organisationen.
- Suchen nach Wörtern, die wie der Suchbegriff klingen.

Sie können Ihre Textsuche in Abfragen von Geschäftsdaten integrieren. Beispielsweise können Sie eine SQL-Abfrage in einer Anwendung codieren, die nach Textdokumenten sucht, die von einem bestimmten Autor innerhalb eines Datumsbereichs erstellt wurden, und die ein bestimmtes Wort oder einen bestimmten Ausdruck enthalten. Unter Verwendung der Text Extender-Programmierschnittstelle können Sie sogar den Benutzern Ihrer Anwendung ermöglichen, die Dokumente anzuzeigen.

Durch die Integration der Volltextsuche in SELECT-Abfragen von DB2 UDB steht Ihnen eine leistungsstarke Abruffunktion zur Verfügung, in der Attribut- und Volltextsuche kombiniert sind. Die folgende SQL-Anweisung zeigt ein Beispiel:

```
SELECT * FROM MyTextTable
  WHERE version = '2'
  AND DB2TX.CONTAINS (
              DB2BOOKS_HANDLE,
              '"authorization"
                IN SAME PARAGRAPH AS "table"
                AND SYNONYM FORM OF "delete"') = 1
```

DB2TX.CONTAINS ist eine von mehreren Text Extender-Suchfunktionen. DB2BOOKS_HANDLE ist der Name einer Kennungsspalte, die auf die Spalte DB2BOOKS verweist, die die zu durchsuchenden Textdokumente enthält. Der Rest der Anweisung ist ein Beispiel für ein Suchargument, das nach `authorization` sucht, das im selben Absatz vorkommen soll wie `table` und `delete` oder einem Synonym von `delete`.

Andere DB2 Extender

Mit den anderen Extendern der Produktfamilie können Sie nach einer Kombination aus Abbild-, Video- und Audiodatentypen *in einer SQL-Abfrage* suchen.

Wie der Text Extender definieren diese Extender neue Datentypen und Funktionen, wobei sie die in DB2 UDB integrierte Unterstützung für benutzerdefinierte Typen und benutzerdefinierte Funktionen verwenden. Sie können eine beliebige Kombination dieser Datentypen, d.&rbl;h. Abbild, Ton und Video, mit einer Suchabfrage nach Text koppeln.

Die Extender nutzen die Unterstützung von DB2 UDB für große Objekte von bis zu 2 Gigabyte und für Auslöser, die eine datenbanktabellenübergreifende Integritätsprüfung bieten, um die referentielle Integrität von Multimedia-Daten sicherzustellen.

Image Extender

Mit dem Image Extender können Ihre Anwendungen folgendes:

- Importieren und Exportieren von Abbildern und deren Attributen in eine und aus einer Datenbank.
- Steuerung des Zugriffs auf Abbilder mit der gleichen Sicherungsstufe wie für herkömmliche Geschäftsdaten.
- Auswählen und Aktualisieren von Abbildern auf der Basis von Abbildattributen, wie z.&rbl;B. Format, Breite und Höhe.
- Anzeigen von Miniabbildern und ganzen Abbildern.

Sie können eine Abbildabfrage in herkömmliche Abfragen von Geschäftsdatenbanken integrieren. Beispielsweise können Sie eine SQL-Anweisung in einer Anweisung programmieren, die die Miniabbilder alle Bilder zurückgeben soll, deren Breite und Höhe kleiner als 512 x 512 Pixel ist und deren Preis niedriger als 500 Dollar ist. Außerdem sollen die Namen der einzelnen Fotografen der Bilder aufgelistet werden. Unter Verwendung des Image Extender können Sie sogar den Benutzern Ihrer Anwendung ermöglichen, die Abbilder anzuzeigen.

Video Extender

Mit dem Video Extender können Ihre Anwendungen folgendes:

- Importieren und Exportieren von Videoclips und deren Attributen in eine und aus einer DB2 UDB-Datenbank.
- Auswählen und Aktualisieren von Videoclips auf der Basis von Videoattributen, z.&rbl;B. Komprimierungsmethode, Länge, Vollbildrate und Anzahl an Vollbildern.
- Abrufen von bestimmten Aufnahmen in einem Videoclip über die Aufnahmeermittlung.
- Wiedergeben von Videoclips.

Sie können eine Videoabfrage in herkömmliche Abfragen von Geschäftsdatenbanken integrieren. Beispielsweise können Sie eine SQL-Anweisung in einer Anwendung codieren, die die Miniabbilder und Namen der Werbeagenturen aller Werbespots zurückgeben soll, deren Länge kleiner als 30 Sekunden ist, deren Vollbildrate größer als 15 Vollbilder pro Sekunde ist und die im Script zum Werbespot Bemerkungen wie etwa “OS/2 Warp” enthalten. Unter Verwendung des Video Extender können Sie sogar den Benutzern Ihrer Anwendung ermöglichen, die Werbespots wiederzugeben.

Andere DB2 Extender

Audio Extender

Mit dem Audio Extender können Ihre Anwendungen folgendes:

- Importieren und Exportieren von Audioclips und deren Attributen in eine und aus einer DB2 UDB-Datenbank.
- Auswählen und Aktualisieren von Audioclips auf der Basis von Audioattributen, z.&rbl;B. Anzahl der Kanäle, Länge und Abtastrate.
- Wiedergeben von Audioclips.

Der Audio Extender unterstützt eine Vielzahl von Audiodateiformaten, wie z. B. WAVE und MIDI. Wie der Video Extender arbeitet der Audio Extender mit verschiedenen Audio-Servern auf Dateibasis.

Mit dem Audio Extender können Ihre Anwendungen Audiodaten und herkömmliche Geschäftsdaten in einer Abfrage integrieren. Beispielsweise können Sie eine SQL-Anweisung in einer Anwendung codieren, die die Miniabbilder von CD-Covers und die Namen der Sänger aller Musiksegmente auf der CD zurückgeben soll, deren Länge unter 1 Minute liegt und die 1990 produziert wurden. Unter Verwendung des Audio Extender können Sie sogar den Benutzern Ihrer Anwendung ermöglichen, die Musiksegmente wiederzugeben.

Der Text Extender in der DB2 Client/Server-Umgebung

Abb. 1 zeigt, wie der Text Extender in der DB2 Client/Server-Umgebung integriert ist.

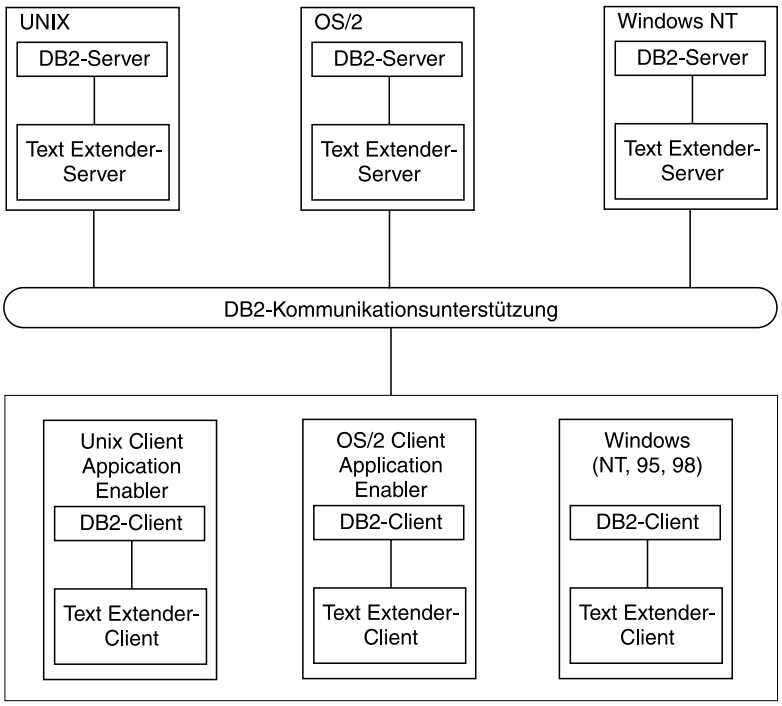


Abbildung 1. Integration des Text Extender in der DB2 Client/Server-Umgebung

! Eine Liste der DB2-Protokolle zur Kommunikationsunterstützung (wie z.&rbl;B. TCP/IP
! oder NETBIOS) für einen Client finden Sie im Buch *DB2 Einstieg* für die entsprechende
! Plattform.

Die Hauptkomponenten des Text Extender sind auf derselben Maschine installiert wie der DB2-Server. Jeweils nur ein Text Extender-Server-Exemplar kann mit einem DB2-Server-Exemplar installiert werden.

Eine Text Extender-Installation ist flexible und kann folgende Komponenten umfassen:

- Einen oder mehrere Text Extender-Server unter einem der Betriebssysteme, die in Abb. 1 gezeigt werden, wobei UNIX die AIX-, SUN-Solaris- und HP-UX-Workstations einschließt.
- AIX-, SUN-Solaris-, HP-UX-, OS/2-, Windows NT-, Windows 98- oder Windows 95-Clients mit Zugriff auf einen oder mehrere ferne Text Extender-Server.
- AIX-Clients, die einen lokalen Server und Zugriff auf ferne Server haben.

Der Text Extender in der DB2 Client/Server-Umgebung

Abb. 2 zeigt eine typische Text Extender-Konfiguration. Um den Text Extender von einem Client aus auszuführen, müssen Sie zunächst einen DB2-Client und einige Text Extender-Dienstprogramme installieren. Diese Dienstprogramme bilden den Text Extender-“Client”, obwohl es kein Client im herkömmlichen Sinn des Wortes ist. Der Client kommuniziert mit dem Server über eine DB2-Client-Verbindung.

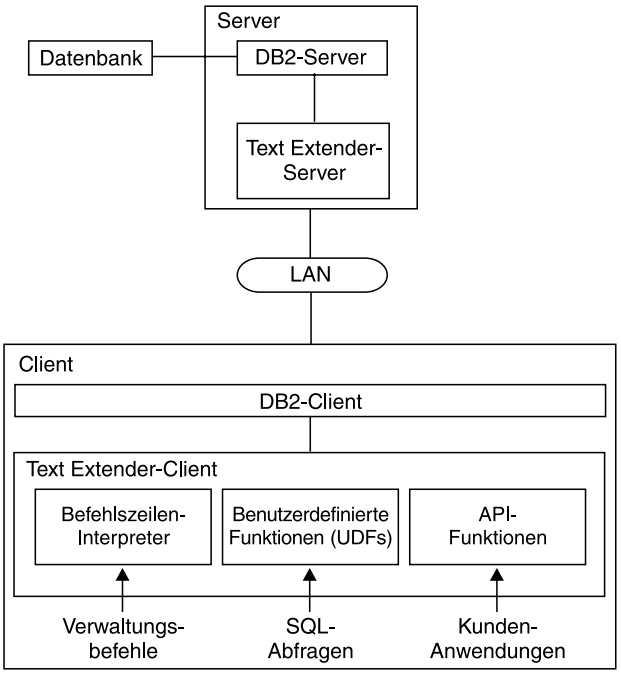


Abbildung 2. Eine Text Extender-Konfiguration

Der Text Extender besteht aus folgenden Hauptkomponenten:

- **Ein Befehlszeilen-Interpreter** Befehle stehen zur Verfügung, mit denen Sie Text in Spalten für die Suche vorbereiten und Textindizes verwalten können.
- **Benutzerdefinierte Funktionen (UDFs)** Funktionen stehen zur Verfügung, die Sie für das Suchen in Text und beispielsweise das Ermitteln der Anzahl an Vorkommen des Suchbegriffs im Text in SQL-Abfragen einschließen können. Zur Verdeutlichung zeigt die Abbildung die UDFs auf dem Client, da sie als Teil einer SQL-Abfrage verwendet werden können. Tatsächlich sind sie Teil der Server-Installation und werden dort ausgeführt. Die UDFs können jedoch von jedem DB2-Client aus ausgeführt werden, ohne daß der Text Extender-Client installiert sein muß.
- **Eine Anwendungsprogrammierschnittstelle (API)** Sie besteht aus Funktionen, die in C-Programmen für die Suche in Text und die Anzeige der Suchergebnisse aufgerufen werden können.

Hinweis

Die Text Extender-Client-Dienstprogramme bieten zwei Hauptfunktionen: Verwaltung und die API. (Die UDFs sind auf dem Server verfügbar.)

Um die Text Extender-Verwaltungsfunktionen oder die API auf dem Client verwenden zu können, müssen Sie den Text Extender-Client installieren. Wenn Sie die Text Extender-Verwaltung nur auf dem Text Extender-Server ausführen, müssen Sie die Text Extender-Client-Dienstprogramme nur auf den Clients installieren, die die API-Funktionen verwenden.

Wenn Sie nur die Suchfunktion auf dem Client unter Verwendung von DB2 UDB-SQL-Anweisung benötigen, müssen Sie den Text Extender-Client nicht installieren. Die gesamte Kommunikation wird von DB2 UDB bearbeitet, und die Text Extender-Suchmaschine wird nur auf dem Server ausgeführt.

Der Text Extender in der partitionierten Datenbankumgebung

Abb. 3 zeigt, wie der Text Extender in der DB2-Umgebung für partitionierte Datenbanken integriert ist.

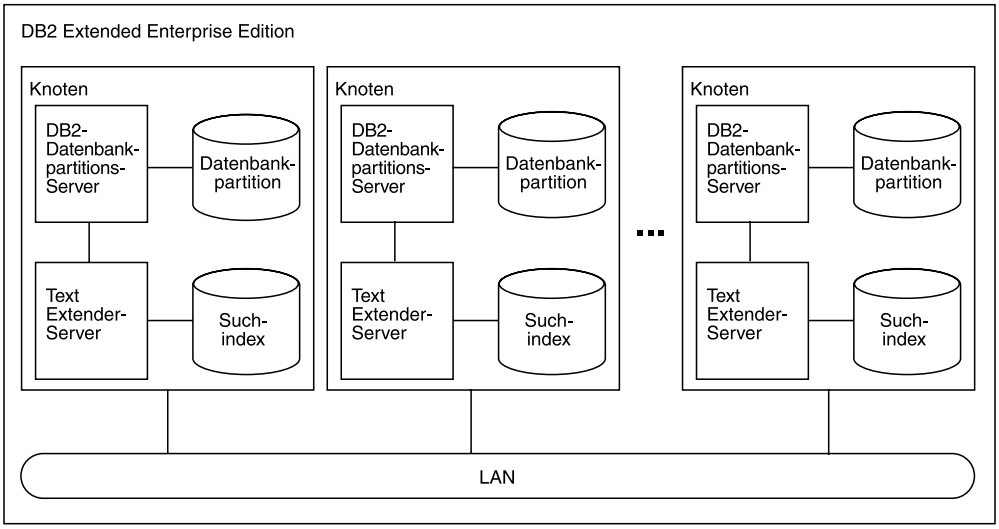


Abbildung 3. Integration des Text Extender in die DB2-Umgebung für partitionierte Datenbanken

Jeder Datenbankpartitions-Server, als *Knoten* bekannt, kann einer separaten Maschine zugeordnet sein, oder sie können sich alle auf einer einzelnen Maschine befinden.

Der Text Extender in der partitionierten Datenbankumgebung

Datenbankpartitions-Server, die sich auf separaten Maschinen befinden, sind als *physische Knoten* bekannt. Datenbankpartitions-Server, die sich auf einer einzelnen Maschine befinden, sind als *logische Knoten* bekannt. Der Text Extender unterstützt beide Konfigurationen.

Eine *Knotengruppe* ist eine benannte Untergruppe von einem oder mehreren Datenbankpartitions-Servern. Wenn eine Text Extender-fähige Tabelle einer Mehrpartitionsknotengruppe zugeordnet ist, befinden sich die Textindizes, die vom Text Extender erstellt werden, auf demselben Knoten wie die entsprechende Tabellenpartition. Hierdurch ist sichergestellt, daß die Daten während der Indexierung nicht zwischen Knoten verschoben werden müssen.

Ein Text Extender-Server wird auf jeder Partition in einer Knotengruppe ausgeführt und verwaltet die indexierten Zeilen der entsprechenden Partition. Sie stehen im Dialog mit dem Text Extender über eine Datenbankpartition. Jede Datenbankpartition kann verwendet werden, um Text Extender-Anforderungen auszugeben. Die Anforderung wird an alle betroffenen Datenbankpartitionen weitergeleitet.

Kapitel 2. Einen Textindex planen

Bevor Sie mit den Schritten beginnen, die in Kapitel 4, „Verwaltung“ auf Seite 43, beschrieben sind, müssen Sie entscheiden, ob Sie einen gemeinsamen Textindex erstellen wollen oder jeweils einzelne Indizes für jede Textspalte. Dieses Kapitel unterstützt Sie bei dieser Entscheidung und berechnet ungefähr, wieviel Plattenspeicherplatz Sie für Textindizes benötigen werden.

Es gibt verschiedene Indextypen, aus denen Sie wählen können: linguistisch, präzise, dual und Ngram. Die Auswahl des Indextyps ist entscheidend. Wenn Sie beispielsweise *linguistisch* als Indextyp auswählen, können Sie nach Wortvariationen und Synonymen des Suchbegriffs suchen. Der Indextyp kann außerdem die Indexierungsleistung und
! die Größe des Indexes beeinflussen. Sie können außerdem die Suchfunktionen von
! mehr als einem Indextyp nutzen, indem Sie mehrere Indizes pro Textspalte erstellen,
! von denen jeder einen anderen Indextyp hat.

Warum Textdokumente indexiert sein müssen

Ein schnelles Informationsabfragesystem durchsucht Textdokumente nicht sequentiell, das würde zu lange dauern. Statt dessen arbeitet es mit einem zuvor erstellten Textindex. Ein Textindex besteht aus signifikanten Begriffen, die aus den Textdokumenten extrahiert wurden, wobei jeder Begriff zusammen mit Informationen zum Dokument, in dem er enthalten ist, gespeichert wird.

Ein Textindex enthält nur relevante Informationen; unwesentliche Wörter, z.&rbl;B. “und”, “von” und “der/die/das”, werden nicht indexiert. Der Text Extender verwendet eine Liste dieser Wörter, bekannt als *Stoppwörter*, um zu verhindern, daß sie indexiert werden. Das Abfragesystem durchsucht den Index nach den angeforderten Begriffen, um zu ermitteln, welche Textdokumente diese Begriffe enthalten.

Hinweis

Wenn Sie die Liste der Stoppwörter ändern müssen, tun Sie dies nur einmal und zwar bei der Installation.

Eine Liste von Stoppwörtern in der jeweiligen Sprache ist in einer Datei gespeichert, die Sie ändern können (siehe „Die Stoppwort- und Abkürzungsdateien ändern“ auf Seite 52). Da aber nur eine Datei für das gesamte System zur Verfügung steht, sollten Sie sie nur einmal ändern, wenn Sie den Text Extender zum ersten Mal einrichten. Wenn Sie die Datei zu einem späteren Zeitpunkt ändern, geben die bestehenden Indizes diese Änderungen nicht wieder.

Gehen wir beispielsweise davon aus, daß einige Dokumente den Namen eines Wochenmagazins, “Jetzt”, enthalten. Wenn Sie dieses Wort aus der Liste der Stoppwörter löschen, wird es indexiert und kann bei zukünftigen Suchen gefunden werden. Alle Indizes, die vor dem Löschen des Stoppwortes erstellt wurden, enthalten jedoch das Wort “jetzt” nicht und eine Suche danach ist nicht erfolgreich.

Warum Textdokumente indexiert sein müssen

Wenn Sie die Liste der Stoppwörter ändern wollen und diese Änderungen im ganzen System gelten sollen, müssen Sie alle Indizes erneut erstellen.

Das Indexieren ist ein aus zwei Schritten bestehender Prozeß. Im ersten Schritt werden die Textdokumente, die indexiert werden müssen, in einer *Protokolltabelle* aufgezeichnet. Dies geschieht jeweils automatisch durch DB2-*Auslöser*, wenn Sie ein Textdokument in einer Spalte einfügen, aktualisieren oder löschen.

Im zweiten Schritt werden die Textdokumente, die in der Protokolltabelle aufgelistet sind, indexiert. Dies kann in regelmäßigen Abständen erfolgen. Die Begriffe der Dokumente, die in der Spalte eingefügt oder geändert wurden, werden zum Index hinzugefügt. Die Begriffe der Dokumente, die aus der Spalte gelöscht wurden, werden aus dem Index entfernt.

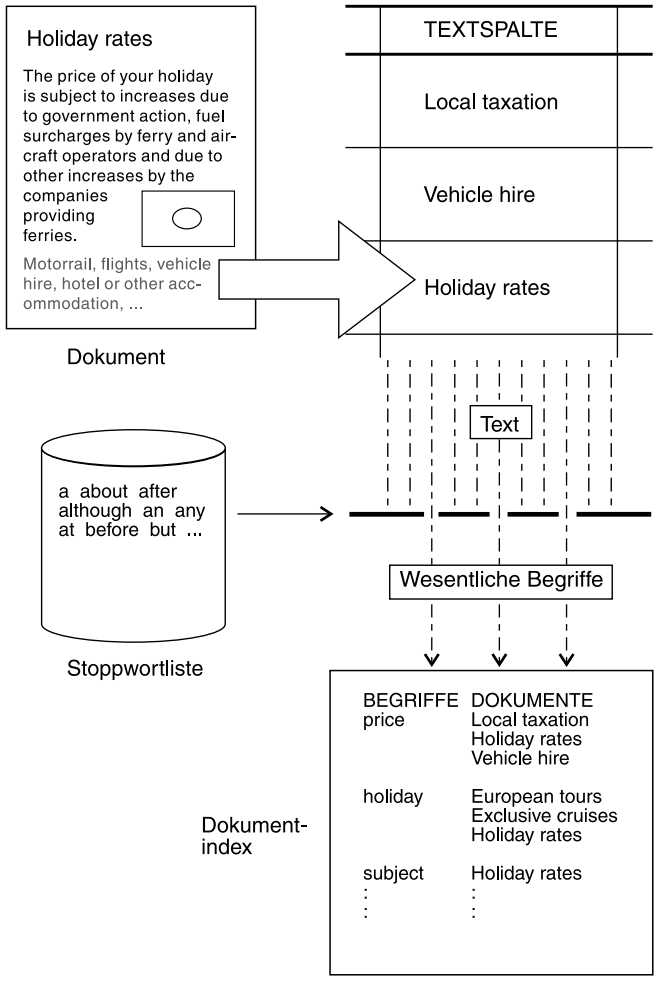


Abbildung 4. Nur wesentliche Begriffe werden indexiert

Indextypen

Sie können einen der folgenden Indextypen einer Spalte zuordnen, die zu suchenden Text enthält: *linguistisch*, *präzise*, *dual* und *Ngram*. Sie müssen entscheiden, welcher Indextyp erstellt werden soll, bevor Sie solche Spalten für die Verwendung durch den Text Extender vorbereiten. Eine detaillierte Beschreibung, wie die einzelnen Indextypen die Verarbeitung auf linguistischer Basis beeinflussen, finden Sie in Kapitel 3, „Verarbeitung auf linguistischer Basis“ auf Seite 19.

Hinweis

Der Text Extender bietet eine breite Palette von Suchoptionen, wobei allerdings nicht alle für alle Indextypen verfügbar sind. Prüfen Sie Tabelle 7 auf Seite 198 und Tabelle 8 auf Seite 199, bevor Sie sich entscheiden, welcher Indextyp verwendet werden soll.

Linguistischer Index

Bei einem linguistischen Index wird die Verarbeitung auf linguistischer Basis angewendet, während der Text der einzelnen Dokumente für das Indexieren analysiert wird. Dies bedeutet, daß Wörter auf ihre Grundform, ihren Stamm, reduziert werden, bevor sie in einem Index gespeichert werden. Der Begriff “Mäuse” wird im Index beispielsweise als `maus` gespeichert.

Bei einer Abfrage anhand eines linguistischen Indexes wird die gleiche Verarbeitung auf linguistischer Basis für die Suchbegriffe angewendet, bevor der Textindex durchsucht wird. Wenn Sie also nach “Mäuse” suchen, wird der Begriff auf seine Grundform `maus` reduziert, bevor die Suche beginnt. Tabelle 1 auf Seite 20 faßt zusammen, wie Begriffe für das Indexieren extrahiert werden, wenn Sie den linguistischen Index verwenden.

Der Vorteil dieses Indextyps liegt darin, daß alle Variationen eines Suchbegriffs mit allen anderen Variationen in einem der indexierten Textdokumente übereinstimmen. Der Suchbegriff `Maus` stimmt mit den Dokumentbegriffen “Maus”, “Mäuse”, “MÄUSE” (Großbuchstaben) usw. überein. Ebenso stimmt der Suchbegriff `Mäuse` mit denselben Dokumentbegriffen überein.

Für diesen Indextyp ist am wenigsten Plattenspeicherplatz erforderlich. Das Indexieren und Suchen kann jedoch länger dauern als für einen präzisen Index.

Die verfügbaren Typen der Verarbeitung auf linguistischer Basis hängen von der Sprache des Dokuments ab. Einzelheiten finden Sie in „Linguistische Funktionen für die unterstützten Sprachen“ auf Seite 30. Folgende Typen stehen zur Verfügung:

- Wort- und Satztrennung
- Verarbeitung bei Satzbeginn
- Entfernung der Silbentrennung
- Normalisieren von Begriffen auf eine Standardform, in der keine Großbuchstaben vorkommen und in der Buchstaben mit Akzenten (Sonderzeichen) wie z.&rbl;B. “ü”

in eine Form ohne Akzente geändert werden. Beispielsweise wird das Wort “Tür” als `tuer` indexiert.

- Reduzieren von Begriffen auf ihre Grundform, ihren Stamm. Beispielsweise wird “gekauft” als `kaufen` und “Mäuse” als `maus` indexiert.

  Hinweis

  Wortfragmente (Worte, die mit Platzhalterzeichen versehen sind), können nicht auf eine Grundform reduziert werden. Wenn Sie also nach `geschwo%` suchen, werden Sie das Wort “geschwommen” nicht finden, da es im Index auf seine Grundform reduziert ist. Um es zu finden, müßten Sie nach `schwi%` suchen.

- Wortdekomposition, wobei Komposita wie “Wetterbericht” nicht nur als `wetterbericht`, sondern auch als `wetter` und `bericht` indexiert werden.
- Stoppwortfiltern, wobei irrelevante Begriffe nicht indexiert werden. “Ein Bericht über alle Tiere” wird indexiert als `bericht` und `tier`.
- Wortartfiltern, das dem Stoppwortfiltern ähnelt: nur Substantive, Verben und Adjektive werden indexiert. “Ich fahre mit meinem Auto schnell” wird indexiert als `fahren` und `auto`. Die Wörter “Ich” und “meinem” werden als Stoppwörter gelöscht, aber zusätzlich wird das Adverb “schnell” durch das Wortartfiltern entfernt.

Präziser Index

In einem präzisen Index werden die Begriffe in einem Textdokument exakt so indexiert, wie sie im Dokument erscheinen. Beispielsweise kann der Suchbegriff `Maus` den Begriff “Maus” finden, aber nicht “Mäuse” und auch nicht “maus”; bei der Suche in einem präzisen Index wird die Groß-/Kleinschreibung beachtet.

In einer Abfrage wird die gleiche Verarbeitungsweise auf die Abfragebegriffe angewendet, die anschließend mit den Begriffen, die im Index gefunden werden, verglichen werden. Dies bedeutet, daß die gefundenen Begriffe exakt mit dem Suchbegriff übereinstimmen. Sie können Platzhalterzeichen verwenden, um die Suche zu erweitern: beispielsweise kann der Suchbegriff `experiment*` die Begriffe “experimental”, “experimentell” usw. finden.

Tabelle 2 auf Seite 20 liefert einige Beispiel, wie Begriffe für die Indexierung aus dem Dokumenttext extrahiert werden, wenn Sie einen präzisen Index verwenden.

Der Vorteil dieses Indextyps liegt darin, daß die Suche präziser und das Indexieren und Abfragen schneller ist. Da jede unterschiedliche Form und Schreibweise jedes Begriffs indexiert wird, wird mehr Plattenspeicherplatz benötigt als für den linguistischen Index.

Folgende linguistische Prozesse werden zum Indexieren von Textdokumenten für einen präzisen Index verwendet:

Wort- und Satztrennung

Stoppwortfiltern

Dualer Index

Ein dualer Index ist eine Kombination aus präzisem und linguistischem Index. Er enthält die normalisierte Form (Standardform nur in Kleinbuchstaben, ohne Akzente, Sonderzeichen), die Grundform (z.&rbl;B. die Infinitive von Verben) und die präzise Form jedes Begriffs.

Bei diesem Indextyp kann der Benutzer für jeden Suchbegriff entscheiden, ob die Suche linguistisch oder präzise erfolgen soll.

In einer Abfrage können Sie die Verarbeitung wählen, die für die Abfragebegriffe angewendet werden soll:

- Linguistische Suche (Option STEMMED FORM OF): jede Form und Groß-/Kleinschreibung eines Suchbegriffs stimmt mit jeder anderen Form und Groß-/Kleinschreibung in einem Textdokument überein.
- Präzise Suche (Option PRECISE FORM OF): nur die exakte Form und exakte Groß-/Kleinschreibung stimmt mit dem Suchbegriff überein.

! Dieser Indextyp benötigt am meisten Plattenspeicherplatz. Das Indexieren und Suchen
! ist langsamer als für einen linguistischen Index. Dieser Indextyp wird für eine große
! Anzahl an Textdokumenten nicht empfohlen.

Hinweis

In einem dualen Index werden Wortfragmente immer so gesucht wie in einem präzisen Index, d.&rbl;h., bei der Suche nach Übereinstimmungen für Wortfragmente wird die Groß-/Kleinschreibung beachtet.

Ngram-Index

Ein Ngram-Index analysiert Text, indem er Zeichengruppen syntaktisch analysiert. Diese Analyse erfolgt nicht auf der Basis eines Wörterverzeichnisses.

Wenn Ihr Text DBCS-Zeichen enthält, müssen Sie den Ngram-Index verwenden. Kein andere Indextyp unterstützt DBCS-Zeichen.

Dieser Indextyp unterstützt die Suche nach grober Übereinstimmung (“fuzzy”), d.&rbl;h., Sie können Zeichenfolgen finden, die dem angegebenen Suchbegriff ähneln. Beispielsweise findet eine Suche nach `Extender` das inkorrekt geschriebene Wort `Extendrrs`. Sie können außerdem einen erforderlichen Grad der Ähnlichkeit angeben.

Anmerkung: Auch wenn Sie die Suche nach grober Übereinstimmung verwenden, müssen die ersten drei Zeichen übereinstimmen.

Um in einem Ngram-Index eine Suche durchzuführen, bei der die Groß-/Kleinschreibung beachtet werden soll, ist es nicht ausreichend, in der Abfrage das Schlüsselwort PRECISE FORM OF anzugeben. Der Grund dafür ist, daß ein Ngram-Index normalerweise nicht zwischen Groß- und Kleinschreibung der indexierten Zeichen unterscheidet. Sie können in einem Ngram-Index jedoch die

Einen oder mehrere Textindizes für eine Tabelle erstellen

Groß-/Kleinschreibung beachten lassen, wenn Sie bei der Erstellung des Indexes die Option CASE_ENABLED angeben. Anschließend können Sie in Ihrer Abfrage das Schlüsselwort PRECISE FORM OF verwenden.

Wenn die Option CASE_ENABLED verwendet wird, benötigt der Index mehr Speicherbereich und die Suchen dauern länger.

! Die SBCS-CCSIDs, die von Ngram-Indizes unterstützt werden, sind 819, 850 und 1252.
! Die DBCS-CCSIDs, die von Ngram-Indizes unterstützt werden, sind 932, 942, 943, 948,
! 949, 950, 954, 964, 970, 1363, 1381, 1383, 1386, 4946 und 5039.

Obwohl der Ngram-Indextyp für das Indexieren von DBCS-Dokumenten konzipiert wurde, kann er auch für SBCS-Dokumente verwendet werden. Er unterstützt jedoch nur TDS-Dokumente.

Beachten Sie außerdem, daß nicht alle Syntaxoptionen für die Suche unterstützt werden. Eine Zusammenfassung der Regeln und Einschränkungen finden Sie in Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

Den Textindextyp ändern

Wenn Sie feststellen, daß der verwendete Indextyp nicht geeignet ist, können Sie ihn ändern. In diesem Fall wird jedoch der bestehende Index gelöscht, ein leerer Index erstellt und werden Einträge für alle Textdokumente in der Spalte zur Protokolltabelle hinzugefügt, damit sie erneut indexiert werden. Den entsprechenden Befehl dazu finden Sie unter „CHANGE INDEX SETTINGS“ auf Seite 129.

Sie können auswählen, daß die Dokumente sofort erneut indexiert werden oder zum Zeitpunkt, zu dem das regelmäßige Indexieren stattfindet.

Einen oder mehrere Textindizes für eine Tabelle erstellen

! Kapitel 4, „Verwaltung“ auf Seite 43 beschreibt, wie Tabellen vorbereitet werden, so
! daß Sie darin nach Text suchen können. Bevor Sie diese Vorbereitung durchführen,
! müssen Sie jedoch entscheiden, ob Sie einen Textindex erstellen, der für alle
! indexierten Textspalten in einer Tabelle gilt, oder ob Sie mehrere Textindizes erstellen,
! jeweils einen für jede indexierte Textspalte.

Separate Textindizes für jede Textspalte (eine Mehrindextabelle) bieten mehr Flexibilität, Sie können für jede Textspalte einen unterschiedlichen Indextyp erstellen. Diese Flexibilität gilt auch für andere Merkmale, die einer Textspalte zugeordnet sind, z.&rbl;B. wann ihr Index regelmäßig aktualisiert wird und in welchem Verzeichnis der Index gespeichert wird. Eine Beschreibung dieser Merkmale finden Sie in „ENABLE TEXT COLUMN“ auf Seite 139. Indexieren kann ein zeit- und ressourcenintensiver Vorgang sein. Bei Verwendung einer Mehrindextabelle können Sie diesen Vorgang über eine bestimmte Zeitspanne verteilen, indem Sie die Spalten zu unterschiedlichen Zeiten indexieren.

! Wenn Sie die Flexibilität, die eine Mehrindextabelle bietet, nicht benötigen, hat ein
! gemeinsamer Index den Vorteil, daß der Text Extender einfacher zu verwalten ist.

Hinweis

Wenn Sie externe Dateien indexieren wollen (siehe „Externe Textdateien aktivieren“ auf Seite 72), muß die zugeordnete Tabelle eine Mehrindextabelle sein.

Die Größe eines Indexes berechnen

! Der Plattenspeicherplatz, den Sie für einen Index benötigen, hängt von der Größe der
! zu indexierenden Daten und dem Datentyp sowie vom Indextyp ab. Textdokumente, die
! mit einem Textverarbeitungsprogramm geschrieben wurden, benötigen weniger Spei-
! cherbereich, da ihr Inhalt mit Steuerzeichen aufgebaut ist. Als Richtlinie gilt: reservieren
! Sie etwa 0,7 mal die Größe des zu indexierenden Dokuments an Plattenspeicherplatz,
! multiplizieren Sie diese Größe anschließend mit 2, um den temporären Speicherplatz
! für das Reorganisieren des Indexes zu reservieren.

! Wenn Sie mehrere große Indizes haben, sollten Sie sie auf separaten Platteneinheiten
! speichern, besonders, wenn Sie gleichzeitigen Zugriff auf die Indizes während der
! Indexaktualisierung oder Suche haben.

Die Größe eines Indexes berechnen

Kapitel 3. Verarbeitung auf linguistischer Basis

Der Text Extender bietet die Verarbeitung auf linguistischer Basis in folgenden Abfragebereichen:

- **Indexierung**. Wenn der Text Extender Dokumente analysiert, um die Begriffe zu extrahieren, die im Textindex gespeichert werden sollen, wird der Text linguistisch verarbeitet, um die richtigen Begriffe für den Index zu extrahieren. Dies dient dazu, die Abfrage so einfach und schnell wie möglich zu machen.
- **Abfrage**. Wenn der Text Extender den Dokumentindex durchsucht, um Dokumente zu finden, die Vorkommen der von Ihnen angegebenen Suchbegriffe enthalten, werden die Suchbegriffe auch linguistisch verarbeitet, um sie mit den indexierten Begriffen abzugleichen.
- **Anzeigen**. Wenn Sie ein Dokument anzeigen, das nach einer Suche gefunden wurde, wird die Verarbeitung auf linguistischer Basis verwendet, um die im Dokument gefundenen Begriffe hervorzuheben.

Verarbeitung auf linguistischer Basis beim Indexieren

Wenn der Text Extender Dokumente indexiert und abruft, führt er eine linguistische Analyse des Textes durch. Wie Sie aus der folgenden Tabelle ersehen können, hängt der Grad an Verarbeitung auf linguistischer Basis vom Indextyp ab. Für Ngram-Indizes wird keine Verarbeitung auf linguistischer Basis angewendet.

Die Verarbeitung auf linguistischer Basis für das Indexieren von Dokumenten besteht aus:

Basistextanalyse

Erkennen von Begriffen (mit Token versehen)

Normalisieren der Begriffe auf eine Standardform

Erkennen von Sätzen

Reduzieren von Begriffen auf ihre Grundform, ihren Stamm

Stoppwortfiltern

Dekomposition (Trennen von Komposita)

Verarbeitung auf linguistischer Basis beim Indexieren

Tabelle 1 zeigt eine Zusammenfassung, wie Begriffe indexiert werden, wenn der Indextyp **linguistisch** ist und keine zusätzlichen Indexmerkmale angefordert wurden.

Tabelle 1. Begriffsextraktion für einen linguistischen Index

Dokumenttext	**Begriff im Index**	**Verarbeitung auf linguistischer Basis**
Maus Käfer	maus kaefer	Basistextanalyse (Normalisierung)
Mäuse geschwommen	maus schwimmen	Reduktion auf die Grundform
systembasiert Wetterbericht	systembasiert, system basis wetterbericht, wetter bericht	Dekomposition
ein Bericht über Tiere	Bericht tier	Stoppwortfiltern. Die Stoppwörter sind: ein, über.

Zum Vergleich zeigt Tabelle 2 eine Zusammenfassung, wie Begriffe indexiert werden, wenn der Indextyp **präzise** ist.

Tabelle 2. Begriffsextraktion für einen präzisen Index

Dokumenttext	**Begriff im Index**	**Verarbeitung auf linguistischer Basis**
Maus Käfer	Maus Käfer	Keine Normalisierung
Mäuse geschwommen	Mäuse geschwommen	Keine Reduktion auf die Grundform
ein Bericht über Tiere	bericht Tiere	Stoppwortfiltern. Die Stoppwörter sind: ein, über.
systembasiert Wetterbericht	systembasiert Wetterbericht	Keine Dekomposition

Verarbeitung auf linguistischer Basis beim Indexieren

Basistextanalyse

Der Text Extender verarbeitet die Basistextanalyse ohne Verwendung eines elektronischen Wörterverzeichnisses.

Begriffe erkennen, die nichtalphanumerische Zeichen enthalten

Beim Indexieren von Dokumenten werden Begriffe erkannt, auch wenn sie nichtalphanumerische Zeichen enthalten, z.&rbl;B. “$14,225.23”, “Extender-Server” und “10/22/90”.

Die folgenden Zeichen werden als Teil eines Begriffs angesehen:

Akzente

Währungssymbole

Nummerntrennzeichen (wie z.&rbl;B. “/” oder “.”)

Das Zeichen “@” in einer E-Mail-Adresse (nur Englisch)

Das Zeichen “+”

Sprachspezifische Regeln werden auch verwendet, um Begriffe zu erkennen, die folgendes enthalten:

- Präfixe mit Akzent in romanischen Sprachen, z.&rbl;B. `l'aventure` in Französisch.
- Nationale Formate für Datum- und Zeitangaben und für Zahlen.
- Alternativen, z. B: `Verantwortlichkeit/Zuständigkeit`, die z.&rbl;B. im Englischen mit dem Zeichen “/” angegeben werden.
- Abschließende Apostrophe in italienischen Wörtern, wie z.&rbl;B. `securita'`. Es ist üblich in italienischem Text, wenn der Zeichensatz keine Zeichen mit Akzenten enthält, das Akzentzeichen *nach* dem Zeichen einzugeben. Beispielsweise wird “à” als “a'” eingegeben.

Begriffe auf eine Standardform normalisieren

Bei der Normalisierung werden Begriffe in Groß-/Kleinschreibung und Begriffe, die Zeichen mit Akzent oder Sonderzeichen enthalten, auf eine Standardform reduziert. Dies wird standardmäßig beim linguistischen Indextyp durchgeführt, bzw. wenn ein dualer Index mit dem Suchparameter STEMMED FORM OF verwendet wird. (In einem präzisen Index bliebt die Groß-/Kleinschreibung von Buchstaben unverändert - bei Suchvorgängen wird die Groß-/Kleinschreibung beachtet.)

Beispielsweise wird der Begriff `Computer` als `computer` indexiert, der Großbuchstabe wird in einen Kleinbuchstaben geändert. Eine Suche nach dem Begriff `computer` findet nicht nur Vorkommen von `computer`, sondern auch von `Computer`. Der Effekt der Normalisierung während der Indexierung liegt darin, daß Begriffe gleich indexiert werden, unabhängig davon, ob im Dokument Groß- oder Kleinbuchstaben verwendet werden.

Verarbeitung auf linguistischer Basis beim Indexieren

Die Normalisierung wird nicht nur beim Indexieren, sondern auch bei der Abfrage verwendet. Großbuchstaben in einem Suchbegriff werden durch Kleinbuchstaben ersetzt, bevor die Suche ausgeführt wird. Wenn der Suchbegriff beispielsweise `Computer` lautet, ist der Begriff, der in der Suche verwendet wird, `computer`.

Zeichen mit Akzent und Sonderzeichen werden auf ähnliche Weise normalisiert. Bei jeder Variation von `école`, z.&rbl;B. `École`, wird `école`, `Ecole` usw. gefunden. Bei `Bürger` wird `buerger` und bei `Maße` wird `masse` gefunden.

Wenn der Suchbegriff Platzhalterzeichenzeichen enthält, wird die Normalisierung durchgeführt, bevor die Platzhalterzeichen verarbeitet werden. Beispiel: aus `Bür_er` wird `buer_er`.

Sätze erkennen

Sie können nach Begriffen suchen, die in demselben Satz auftreten. Um dies zu ermöglichen, wird jedes Dokument während der Indexierung analysiert, um zu ermitteln, an welcher Stelle jeder Satz beginnt und endet. Das Ende eines Satzes ist durch einen Punkt, ein Ausrufezeichen oder ein Fragezeichen gekennzeichnet, auf den/das ein Leerzeichen folgt. Viele Abkürzungen, die mit einem Punkt enden, werden ignoriert.

Begriffe auf ihre Grundform reduzieren

In einem linguistischen oder dualen Index können Sie beispielsweise nach `Maus` suchen und `Mäuse` finden. Die Begriffe werden beim Indexieren auf ihre Grundform reduziert, der Begriff `Mäuse` wird als `maus` indexiert. Später, bei Verwendung des Suchbegriffs `Maus` wird das Dokument gefunden. Das Dokument wird aber auch gefunden, wenn Sie nach `Mäuse` suchen.

Der Effekt ist, daß Sie Dokumente finden, die Informationen zu Mäusen enthalten, unabhängig davon, welche Variation des Begriffs `Maus` im Dokument auftritt bzw. als Suchbegriff verwendet wird.

Auf die gleiche Weise werden konjugierte Verben auf ihren Infinitiv reduziert, der Begriff `gekauft` wird beispielsweise zu `kaufen`.

Verarbeitung auf linguistischer Basis beim Indexieren

Stoppwortfiltern

Stoppwörter sind Wörter wie Präpositionen oder Pronomen, die sehr oft in Dokumenten auftreten und daher als Suchbegriffe nicht geeignet sind. Solche Wörter werden in einer Stoppwortliste gesammelt, die jedem Wörterverzeichnis zugeordnet ist, und werden beim Indexierungsprozeß ausgeschlossen.

Bei der Stoppwortverarbeitung wird die Groß-/Kleinschreibung nicht beachtet. Ein Stoppwort `über` wird also auch als erstes Wort im Satz, `Über`, ausgeschlossen. Dies gilt normalerweise nicht für einen Ngram-Index, bei dem die Groß-/Kleinschreibung nicht beachtet wird, es sei denn, er wird unter Verwendung einer Option erstellt, bei der die Groß-/Kleinschreibung beachtet wird. Die Stoppwortliste, die in verschiedenen Sprachen geliefert wird, kann geändert werden.

Dekomposition (Trennen von Komposita)

In germanischen Sprachen, wie Deutsch oder Niederländisch, werden viele Komposita verwendet, wie z.&rbl;B. `Versandetiketten`, das sich aus `Versand` und `Etiketten` zusammensetzt. Solche Komposita können in ihre Komponenten aufgeteilt werden.

Für einen präzisen Index werden Komposita unverändert als ein Wort indexiert. Für einen linguistischen oder dualen Index werden Komposita während der Indexierung getrennt. Bei der Suche werden Komposita getrennt, wenn Sie einen linguistischen Index haben oder wenn Sie den dualen Index mit der Option STEMMED FORM OF verwenden.

Die Komponenten werden gefunden, wenn sie in beliebiger Reihenfolge in einem Dokument auftreten, so lange sie innerhalb eines Satzes vorkommen. Wenn Sie beispielsweise nach `Wetterbericht` suchen, wird auch ein Dokument gefunden, das die Phrase `Bericht über das Wetter` enthält.

Der Versuch, einen Begriff zu trennen, wird unternommen, wenn

- Die Sprache des Begriffs Komposita verwendet
- Der Begriff eine bestimmte Mindestlänge hat
- Der Begriff selbst nicht ein Eintrag im elektronischen Wörterverzeichnis ist - Komposita, die allgemein verwendet werden, wie z.&rbl;B. `Geschäftsbericht`, *sind* im deutschen Wörterverzeichnis enthalten.

Wenn eine mögliche Trennung gefunden wird, werden die Komponenten auf ihre Grundform reduziert. Hier einige Beispiele in Dänisch, Deutsch und Niederländisch.

Verarbeitung auf linguistischer Basis für die Abfrage

Kompositum	Komponenten
børsmæglerselskab	børsmæglerselskab børs mægler selskab
Kindersprachen	kindersprache kind sprache
probleemkinderen	probleemkinderen probleemkind kind probleem

Verarbeitung auf linguistischer Basis für die Abfrage

Die Abfrageverarbeitung zielt darauf ab, Suchbegriffe zu "schwächen", so daß die Abrufrate bei Suchen erhöht wird, d.&rbl;h., mehr relevante Dokumente gefunden werden. Zwei Basisoperationen können für Abfragebegriffe angewendet werden, um dieses Ziel zu erreichen: Erweiterungen und Reduktionen. Darüber hinaus beinhalten einige Suchbegriffoperationen sowohl Erweiterung als auch Reduktion.

- Bei Erweiterungen wird ein Wort oder ein Mehrwortbegriff aus dem Suchbegriff genommen und einer Gruppe von alternativen Suchbegriffen zugeordnet, die ihrerseits Mehrwortbegriffe sein können. Der Quellenausdruck und die Gruppe von Zielbegriffen bilden einen Boolesche OR-Ausdruck in der Abfragesprache des Text Extender. Da bei Erweiterungen der Quellenbegriff unverändert bleibt, sind Erweiterungen zu einem gewissen Grad unabhängig vom Indextyp. Zu Erweiterungsoperationen gehören die folgenden:

  Synonymerweiterung

  Thesaurus-Erweiterung

- Bei Reduktionen wird der Suchbegriff in eine Form geändert, die allgemeiner ist als die vom Benutzer angegebene. Da hierbei der Suchbegriff geändert wird, sind Reduktionen abhängig vom Indextyp, um sicherzustellen, daß der geänderte Begriff übereinstimmt. Daher leitet der Text Extender Reduktionsinformationen aus dem Indextyp oder den Indextypen ab, für den/die die Abfrage gestellt ist. Zu Reduktionen gehören die folgenden:

  Reduktion auf Grundform (siehe „Begriffe auf ihre Grundform reduzieren“ auf Seite 22)
  Normalisierung
  Stoppwörter

- Einige Operationen ändern nicht nur den Suchbegriff, sondern erweitern ihn auch mit einer Gruppe von alternativen Begriffen. Aufgrund der zugehörigen Reduktion hängen diese Operationen erneut von Informationen ab, die im Index enthalten sind. Zu diesen Operationen gehören die folgenden:

  Platzhalterzeichen für Zeichen und Wörter
  Lauterweiterung

Synonyme

Synonyme sind semantisch zugehörige Wörter. Normalerweise haben diese Wörter dieselbe Wortklasse(n) (z.&rbl;B. Substantiv, Verb, etc.) wie der Quellenbegriff. Synonyme werden von einer separaten Datei für jede Sprache abgerufen. Sie werden immer in der Grundform zurückgegeben und sind, mit wenigen Ausnahmen, keine Mehrwortbegriffe. Suchbegriffe werden immer auf ihre Grundform reduziert, wenn nach Synonymen gesucht wird. Hier sind einige Beispiele für die Synonyme eines Wortes in drei Sprachen:

- Englisch

  word:
  comment remark statement utterance term expression
  communication message assurance guarantee warrant bidding command
  charge commandment dictate direction directive injunction instruction
  mandate order news advice intelligence tidings gossip buzz cry
  hearsay murmur report rumor scuttlebutt tattle tittle-tattle
  whispering

- Französisch

  mot:
  expression parole terme vocable lettre billet missive epître
  plaisanterie

- Deutsch

  Wort:
  Vokabel Bezeichnung Benennung Ausdruck Begriff Terminus
  Ehrenwort Brocken Bekräftigung Versprechen Zusicherung Gelöbnis
  Beteuerung Manneswort Schwur Eid Ausspruch

Verarbeitung auf linguistischer Basis für die Abfrage

Thesaurus-Erweiterung

Ein Suchbegriff kann unter Verwendung von Thesaurus-Begriffen erweitert werden, die über eine spezifische Relation erreicht werden. Diese Relationen können hierarchisch (z.&rbl;B. die Relation “Narrower Term” (engerer Begriff)), oder assoziativ (z.&rbl;B. eine Relation “Related Term” (zugehöriger Begriff)) sein oder es kann sich um eine Synonymrelation handeln. Ein Thesaurus-Begriff kann ein Mehrwortbegriff sein und ist dies auch häufig.

In „Thesaurus-Konzepte“ auf Seite 32 wird die Thesaurus-Erweiterung detaillierter beschrieben.

Lauterweiterung

Bei der Lauterweiterung werden einzelne Wörter durch eine Gruppe von ähnlich klingenden Wörtern erweitert. Dies ist besonders nützlich, wenn die exakte Schreibweise des Suchbegriffs nicht bekannt ist.

Platzhalterzeichen für Zeichen und Wörter

Die Verwendung von Platzhalterzeichen ist eine nicht linguistische Erweiterungstechnik, bei der ein regulärer Ausdruck durch die Disjunktion aller indexierten Wörter ersetzt wird, die diesen Ausdruck erfüllen. Weder für einen Ausdruck mit Platzhalterzeichen noch für eine seiner Erweiterungen kann die Reduktion auf die Grundform, die Stoppwortextraktion oder eine andere Erweiterungstechnik angewendet werden. Dies hat möglicherweise den Effekt, daß beispielsweise eine Form eines unregelmäßigen Verbs, wie etwa `geschwommen`, bei der Suche nach dem Begriff mit Platzhalterzeichen `geschwo*` in einem präzisen Index eine Übereinstimmung findet, nicht aber in einem linguistischen Index, in dem diese Form auf die Grundform `schwimmen` reduziert wurde.

! Wenn Sie Platzhalterzeichen für Wörter verwenden, kann die Leistung sehr niedrig
! werden, besonders, wenn Sie in großen Indizes suchen.

Verarbeitung auf linguistischer Basis für die Anzeige

Die Verarbeitung auf linguistischer Basis wird auch verwendet, wenn Sie nach einer Suche Dokumente anzeigen, die gefunden wurden. Dies wird in zwei Stufen durchgeführt:

1. Basistextanalyse: Normalisierung und Begriffserweiterung
2. Erweiterte Übereinstimmung

Stufe 1: Normalisierung und Begriffserweiterung

Die erste Stufe wird ohne Verwendung eines elektronischen Wörterverzeichnisses durchgeführt.

Normalisierung

Die Normalisierung wird in „Basistextanalyse“ auf Seite 21 beschrieben.

Begriffserweiterung

Die Begriffserweiterung ist die Umkehrung der Reduzierung eines Begriffs auf seine Grundform. Wenn der Index linguistisch ist bzw. wenn der Index dual ist und der Suchbegriff die Option STEMMED FORM OF für die Verarbeitung auf linguistischer Basis enthält, werden die Suchbegriffe auf ihre Grundform reduziert, bevor die Suche beginnt.

Ebenso werden, wenn Sie einen linguistischen oder dualen Index haben, die Begriffe eines Dokuments auf ihre Grundform reduziert, bevor sie zum Index hinzugefügt werden. Dokumente werden daher auf der Basis der Grundform eines Begriffs gefunden.

Wenn Sie ein gefundenes Dokument anzeigen, erwarten Sie jedoch, daß alle Varianten der Grundform hervorgehoben werden. Um diese Varianten hervorzuheben, wird der gefundene Grundformbegriff erweitert.

Alle Varianten (Flexionsformen) für jeden gefundenen Begriff in den Wörterverzeichnissen können erstellt werden. Dies sind die Flexionsformen, die für das deutsche Wort `gehen` erstellt werden:

```
gegangen  geh    gehe    gehen    gehend  gehest  gehet  gehst
ging      ginge  gingen  gingest  ginget  gingst  gingt  geht
```

Verarbeitung auf linguistischer Basis für die Anzeige

Stufe 2: Erweiterte Übereinstimmung

Die zweite Stufe ist die erweiterte Übereinstimmung, die in den seltenen Fällen verwendet werden kann, in denen mit der Basistextanalyse und der Normalisierung ein gefundener Begriff nicht hervorgehoben werden kann. Die erweiterte Übereinstimmung findet die eher abgeschwächten Übereinstimmungen.

Sie wählen die erweiterte Übereinstimmung aus, indem Sie DES_EXTENDED als Parameter in der API-Funktion DesOpenDocument angeben.

Die erweiterte Übereinstimmung verwendet die gleiche Verarbeitung auf linguistischer Basis, die während der Erstellung eines linguistischen Indexes durchgeführt wird.

Dies sind die Fälle, in denen durch die erweiterte Übereinstimmung zusätzliche Übereinstimmungen gefunden werden:

- Wenn der Suchbegriff Platzhalterzeichen enthält und eine Flexionsform ist.

  Die Platzhalterzeichen werden verarbeitet, die Reduktion auf den Wortstamm wird für den Suchbegriff durchgeführt und die entsprechenden Dokumente werden gefunden. Ohne die erweiterte Übereinstimmung würde Text, der mit den angegebenen Suchkriterien übereinstimmt, nicht hervorgehoben werden.

  Beispiel: Ein Dokument enthält den flektierten Begriff `schwamm`.

  – Während der Indexierung wird dieser Begriff auf `schwimmen` reduziert.

  – Wenn der Suchbegriff `schwi%` lautet, wird das obige Dokument gefunden, da die Reduktion auf den Wortstamm `schwimmen` ergibt.

  – Ohne die erweiterte Übereinstimmung werden nur die Wörter hervorgehoben, die mit dem Begriff `schwi%` übereinstimmen. Mit der erweiterten Übereinstimmung wird auch der flektierte Begriff `schwamm` hervorgehoben.

- Wenn Komposita indexiert wurden.

  Wenn ein Dokument in deutscher Sprache ein Kompositum enthält und unter Verwendung eines linguistischen Indexes indexiert wird, enthält der Index des Dokuments die Komponenten des Kompositums und das Kompositum selbst. Wenn Sie nach einer Komponente eines Kompositums suchen, werden die Dokumente gefunden, die das Kompositum enthalten, aber ohne die erweiterte Übereinstimmung wird das Wort nicht hervorgehoben.

  Beispiel: Ein Dokument enthält das deutsche Wort `Apfelbaum`.

  – Während der Indexierung in einen linguistischen Index wird dieses Wort auf `apfel` und `baum` reduziert.

  – Wenn der Index nach dem Begriff `baum` durchsucht wird, wird der Begriff `Baum` gefunden und das Dokument, das diesen Begriff enthält, über den Index ermittelt.

  – Ohne die erweiterte Übereinstimmung werden keine Begriffe hervorgehoben, da das Dokument `Apfelbaum` enthält, aber nicht `Baum`. Mit der erweiterten Übereinstimmung wird das Kompositum `Apfelbaum` getrennt und die Komponente `Baum` wird für die Hervorhebung gefunden.

Verarbeitung auf linguistischer Basis für die Anzeige

- Wenn Wörter am Ende einer Zeile getrennt werden.

  Wenn ein Silbentrennungsstrich automatisch durch ein Textverarbeitungsprogramm eingefügt wird, kann das getrennte Wort gefunden und hervorgehoben werden. Wenn der Silbentrennungsstrich jedoch durch den Benutzer eingegeben wurde, werden die Dokumente, die das Wort enthalten, zwar gefunden, aber ohne die erweiterte Übereinstimmung wird das Wort nicht hervorgehoben.

  Beispiel: Ein Dokument enthält das getrennte englische Wort `container`, das wie folgt am Ende einer Zeile umbrochen wird:

  ```
  Another name for a folder is a con-
  tainer.
  ```

  – Während der Indexierung wird das Wort zu `container` normalisiert.

  – Wenn der Index nach dem Begriff `container` durchsucht wird, werden der Begriff und das Dokument, das ihn enthält, gefunden.

  – Ein Versuch wird unternommen, alle Wörter im Dokument, die mit `container` übereinstimmen, hervorzuheben. Ohne die erweiterte Übereinstimmung wird eine Übereinstimmung nur gefunden, wenn der Silbentrennungsstrich in `con-tainer` von einem Textverarbeitungsprogramm eingefügt und nicht vom Benutzer eingegeben wurde.

Linguistische Funktionen für die unterstützten Sprachen

Linguistische Funktionen für die unterstützten Sprachen

Die folgende Tabelle zeigt, wie Sprachen linguistisch vom Text Extender unterstützt werden. Die Verarbeitung von Eigennamen, Abkürzungen und Domänenbegriffen ist nur für amerikanisches und britisches Englisch verfügbar.

Tabelle 3. Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden

Sprache	**Linguistische Funktionen**
Brasilianisch	Keine Dekomposition
Französisch (Kanada)	Keine Dekomposition
Katalanisch	Keine Dekomposition
Dänisch	Alle Funktionen
Niederländisch	Alle Funktionen
Finnisch	Basistextanalyse, Normalisierung, Entfernung der Silbentrennung, Stoppwortfilter
Französisch	Keine Dekomposition
Deutsch	Alle Funktionen
Isländisch	Reduktion auf Grundform, Entfernung der Silbentrennung, Stoppwortfilter und Wortartfilter werden nicht unterstützt
Italienisch	Keine Dekomposition
Norwegisch (Bokmal)	Alle Funktionen
Norwegisch (Nynorsk)	Keine Synonyme
Norwegisch (Bokmal und Nynorsk)	Keine Synonyme
Portugiesisch	Keine Dekomposition
Spanisch	Keine Dekomposition
Schwedisch	Alle Funktionen
Deutsch (Schweiz)	Alle Funktionen
Britisches Englisch	Keine Dekomposition
Amerikanisches Englisch	Keine Dekomposition

Wörterverzeichnisse, Stoppwortlisten und Abkürzungslisten

! Die folgende Tabelle zeigt die unterstützten Sprachen und die Namen der Dateien, die
! als Wörterverzeichnisse, Stoppwortlisten und Abkürzungslisten geliefert werden. Die
! Wörterverzeichnisdateien liegen im Binärformat vor und können nicht geändert werden.
! Die Stoppwortlisten und die Abkürzungslisten (falls vorhanden) liegen im
! unstrukturierten Dateiformat vor und können geändert werden. Wenn Sie eine dieser
! Dateien ändern, stellen Sie sicher, daß Sie die jeweilige Codepage für die Sprache ver-
! wenden.

Die Dateien werden durch ihre Erweiterung voneinander unterschieden:

Inhalt	Erweiterung
Wörterverzeichnis	DIC
Stoppwortliste	STW
Abkürzungsliste	ABR

Tabelle 4 (Teil 1 von 2). Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden

Sprache	Dateiname	Wörter-ver-zeichnis	Stopp-wort-liste	Abkür-zungsliste	Code-page
! Arabisch	arabic	X	X		864
Portugiesisch (Brasilien)	brazil	X	X		850
Französisch (Kanada)	canadien	X	X	X	850
Katalanisch	catala	X	X		850
Dänisch	dansk	X	X	X	850
Niederländisch	nederlnd	X	X	X	850
Finnisch	suomi	X	X		850
Französisch	francais	X	X	X	850
Deutsch	deutsch	X	X	X	850
! Hebräisch	hebrew	X	X	X	862
Isländisch	islensk	X	X		850
Italienisch	italiano	X	X	X	850
Norwegisch (Bokmal)	norbook	X	X	X	850

Tabelle 4 (Teil 2 von 2). Linguistische Funktionen, die für die verschiedenen Sprachen verwendet werden

Sprache	Dateiname	Wörter-ver-zeichnis	Stopp-wort-liste	Abkür-zungsliste	Code-page
Norwegisch (Nynorsk)	norntn	X	X		850
Portugiesisch	portugal	X	X		850
! Russisch	russian	X	X	X	866
Spanisch	espana	X	X	X	850
Schwedisch	svensk	X	X	X	850
Deutsch (Schweiz)	dschweiz	X	X	X	850
Britisches Englisch	uk	X	X	X	850
Amerikanisches Englisch	us	X	X	X	850

Thesaurus-Konzepte

Ein Thesaurus ist ein gesteuertes Wörterverzeichnis von semantisch verbundenen Begriffen, das normalerweise einen bestimmten Themenbereich abdeckt. Es kann als ein semantisches Netz betrachtet werden, in dem jeder Begriff durch einen Knoten dargestellt ist. Wenn zwei Begriffe in Relation zueinander stehen, sind ihre Knoten durch eine Verbindung verknüpft, die mit dem Relationsnamen gekennzeichnet ist. Alle Begriffe, die direkt in Relation zu einem bestimmten Begriff stehen, können erreicht werden, indem alle Verbindungen, die von dessen Knoten ausgehen, verfolgt werden. Weitere zugehörige Begriffe können erreicht werden, indem iterativ alle Verbindungen verfolgt werden, die von den Knoten ausgehen, die im vorherigen Schritt erreicht wurden. Abb. 5 auf Seite 33 zeigt ein Beispiel der Struktur eines sehr kleinen Thesaurus.

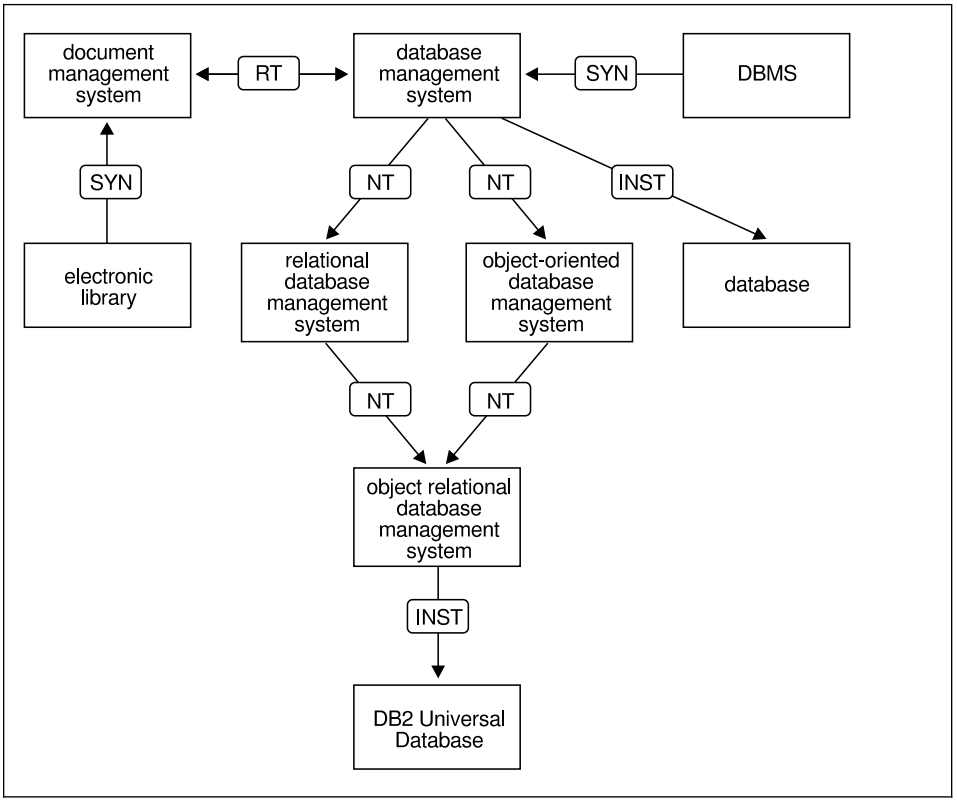


Abbildung 5. Ein als Netz angezeigter Thesaurus

Mit dem Text Extender können Sie einen Suchbegriff erweitern, indem Sie zusätzliche Begriffe aus einem Thesaurus, den Sie zuvor erstellt haben, hinzufügen. In Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, finden Sie Informationen, wie die Thesaurus-Erweiterung in einer Abfrage verwendet wird.

Um einen Thesaurus für die Verwendung in einer Suchanwendung erstellen zu können, ist eine Thesaurus-Definitionsdatei erforderlich, die in ein internes Format, das Thesaurus-Wörterverzeichnis, kompiliert werden muß.

! Das Wörterverzeichnisformat, das von einem linguistischen und einem präzisen Index
! verwendet wird, unterscheidet sich vom Format, das von einem Ngram-Index verwendet
! wird. Daher werden mit dem Produkt zwei unterschiedliche Thesaurus-Compiler gelie-
! fert. Sie unterscheiden sich nicht nur in den zugrundeliegenden Konzepten, sondern es
! sind auch unterschiedliche Quellenformate erforderlich. Sie sollten zunächst ent-
! scheiden, welchen Indextyp Sie verwenden wollen, bevor Sie beginnen, die Thesauri
! für Ihre Anwendungen zu definieren.

Die Basiskomponenten eines Thesaurus sind “Begriffe” und “Relationen”.

Thesaurus-Konzepte

Begriffe

Ein Begriff ist ein Wort oder ein Ausdruck, das/der ein Konzept innerhalb der Themendomäne des Thesaurus angibt. Beispielsweise könnten die folgenden Begriffe in einem oder mehreren Thesauri stehen:

```
data processing (Datenverarbeitung)
helicopter (Hubschrauber)
gross national product (Bruttosozialprodukt)
```

In einem Text Extender-Thesaurus werden Begriffe entweder als Deskriptoren oder Nicht-Deskriptoren klassifiziert. Ein *Deskriptor* ist ein Begriff in einer Klasse von Synonymen, der der bevorzugte Begriff für die Indexierung und Suche ist. Die anderen Begriffe in der Klasse werden als *Nicht-Deskriptoren* bezeichnet. Beispielsweise sind `Umriß` und `Form` Synonyme, wobei `Form` der Deskriptor und `Umriß` ein Nicht-Deskriptor sein könnte.

Ein Ngram-Thesaurus unterscheidet nicht zwischen Deskriptoren und Nicht-Deskriptoren.

Relationen

Eine Relation ist ein Ausdruck einer Beziehung zwischen zwei Begriffen. Relationen haben die folgenden Merkmale:

- Die *Tiefe* einer Relation ist die Anzahl an Ebenen, über die sich die Relation erstreckt. Sie wird in der Suchsyntax unter Verwendung des Schlüsselworts THESDEPTH angegeben. In Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, finden Sie Informationen, wie die Thesaurus-Erweiterung in einer Abfrage verwendet wird.
- Die *Direktionalität* einer Relation gibt an, ob die Relation zwischen zwei Begriffen wirklich gleich ist (bidirektional) oder nur aus einer Richtung besteht (unidirektional).

Die Thesaurus-Erweiterung kann jede Relation verwenden, die im Thesaurus definiert ist. Sie können auch die Tiefe der Erweiterung angeben. Dies ist die maximale Anzahl an Übergängen von einem Quellenbegriff zu einem Zielbegriff. Beachten Sie jedoch, daß die Begriffsgruppe möglicherweise exponential wächst, wenn die Tiefe erhöht wird.

Das folgende Beispiel für den englischen Begriff 'health' (Gesundheit) zeigt die Begriffe, die neu hinzugefügt werden, wenn die Tiefe zunimmt.

```
health

health service, paramedical, medicine, illness

allergology, virology, veterinary medicine, toxicology, surgery,
stomatology, rhumatology, radiotherapy, psychiatry, preventive
medicine, pathology, odontology, nutrition, nuclear medicine,
neurology, nephrology, medical check up, industrial medicine,
hematology, general medicine, epidemiology, clinical trial,
cardiology, cancerology
```

Text Extender-Thesaurus-Relationen

Dies sind die Relationsarten, die von einem Text Extender-Thesaurus geboten werden:

- Assoziativ
- Synonym
- Hierarchisch
- OTHER

In einem Text Extender-Thesaurus sind keine vordefinierten Relationen vorhanden. Sie können jeder Relation einen Namen geben, z.&rbl;B. BROADER TERM, für den ein Kurzzeichen, wie beispielsweise BT, definiert sein kann. Die allgemeinen Relationen, die im Thesaurus-Design verwendet werden, sind:

- BT oder BROADER TERM (weiterer Begriff)
- NT oder NARROWER TERM (engerer Begriff)
- RT oder RELATED TERM (zugehöriger Begriff)
- SYN oder SYNONYM
- USE
- UF oder USE FOR

Assoziativ: Eine assoziative Relation ist eine bidirektionale Relation zwischen Deskriptoren, die auf beliebige Tiefe erweitert werden kann. Sie bindet zwei Begriffe, die weder äquivalent noch hierarchisch sind, dennoch semantisch so eng verbunden, daß die Verbindung zwischen ihnen möglicherweise zusätzliche Begriffe für die Verwendung beim Indexieren oder Abrufen anbietet.

Assoziative Relationen werden im allgemeinen als RT (zugehöriger Begriff) entworfen. Beispiele:

hund RT sicherheit
haustier RT tierarzt

Synonym: Wenn eine Unterscheidung zwischen Deskriptoren und Nicht-Deskriptoren vorgenommen wird wie in einem Text Extender-Thesaurus, ist die Synonymrelation zwischen zwei Begriffen, die die gleiche oder eine ähnliche Bedeutung haben, unidirektional. In einer Klasse von Synonymen ist einer der Begriffe als Deskriptor gekennzeichnet. Die anderen Begriffe werden dann als Nicht-Deskriptoren bezeichnet. In „Ngram-Thesaurus-Relation“ auf Seite 36 finden Sie eine Definition der Synonymrelation, wenn keine Unterscheidung zwischen Deskriptoren und Nicht-Deskriptoren vorgenommen wird.

Die allgemeine Kennzeichnung USE führt von einem bestimmten Nicht-Deskriptor zu seinem Deskriptor. Die allgemeine Kennzeichnung USE FOR führt vom Deskriptor zu jedem Nicht-Deskriptor.

katzentier USE katze
rechtsanwalt UF verteidiger

Thesaurus-Konzepte

Hierarchisch: Eine hierarchische Relation ist eine unidirektionale Relation zwischen Deskriptoren, die angibt, daß einer der Begriffe spezifischer oder weniger allgemein als der andere ist. Diese Differenzierung führt zu einer Darstellung der Begriffe als Hierarchie, in der ein Begriff eine Klasse darstellt und sich untergeordnete Begriffe auf die jeweilige Member-Komponente beziehen. Beispielsweise gehört der Begriff “maus” zur Klasse “nagetier”.

BROADER TERM und NARROWER TERM sind hierarchische Relationen. Beispiel:

auto NT limousine
raubkatzen BT leopard

Andere (Other): Eine Relation vom Typ *OTHER* ist die allgemeinste. Sie stellt eine Zuordnung dar, die nicht einfach in eine der anderen Kategorien fällt. Eine Relation vom Typ *OTHER* kann bidirektional oder unidirektional sein, es gibt keine Einschränkungen der Tiefe und Relationen können zwischen Deskriptoren und Nicht-Deskriptoren bestehen.

Diese Relation wird oft für neue Begriffe in einem Thesaurus verwendet, bis die eigentliche Relation mit anderen Begriffen bestimmt werden kann.

Natürlich können Sie Ihre eigene bidirektionale Synonymrelation unter Verwendung der Relationsart *assoziativ* für eine Synonymrelation zwischen Deskriptoren definieren oder sogar unter Verwendung der Relationsart *OTHER* für eine Synonymrelation zwischen beliebigen Begriffen.

Ngram-Thesaurus-Relation

Ein Ngram-Thesaurus unterstützt die folgenden zwei Relationsarten:

- Assoziativ
- Synonym

Zwei vordefinierte Relationen sind vorhanden, jede davon auf der Basis einer dieser zwei Arten. Sie können Ihre eigenen Relationen auf der Basis der Relationsart *assoziativ* definieren. Einzelheiten dazu finden Sie in „Einen Ngram-Thesaurus erstellen“ auf Seite 41.

Assoziativ

Eine assoziative Relation ist eine bidirektionale Relation zwischen zwei Begriffen, die nicht dasselbe Konzept ausdrücken, aber zueinander in Beziehung stehen. Die vordefinierte Relation *RELATED_TO* und alle benutzerdefinierten Relationen basieren auf dieser Relationsart.

Beispiele:

tennis RELATED_TO schläger
deutsch RELATED_TO wurst

Synonym

Eine Synonymrelation ist eine bidirektionale Relation zwischen zwei Begriffen, die dieselbe oder eine ähnliche Bedeutung haben und als gegenseitig als Alternative verwendet werden können. Diese Relation kann beispielsweise für einen Begriff und seine Abkürzung verwendet werden. Die vordefinierte Relation SYNONYM_OF ist die einzige Relation, die auf dieser Relationsart basiert.

Beispiele:

flagge SYNONYM_OF fahne
US SYNONYM_OF United States

Einen Thesaurus erstellen

Siehe auch „Einen Ngram-Thesaurus erstellen“ auf Seite 41.

! Ein englisches Beispiel für eine Thesaurus-Compiler-Eingabedatei, `desthes.sgm`, ist in
! Verzeichnis SAMPLES des Installationspfades gespeichert. Das Verzeichnis für das
! Wörterverzeichnis auf OS/2- und Windows-Systemen ist

! *laufwerkbuchstabe*:\dmb\db2tx\samples

! Auf AIX-, HP-UX- und SUN-Solaris-Systemen ist das Verzeichnis für das Wörterver-
! zeichnis

! *DB2TX_INSTOWNER* /db2tx/samples

! Eine kompilierte Version dieses Thesaurus und seiner SGML-Eingabedatei wird im Ver-
! zeichnis für das Wörterverzeichnis gespeichert.

! *laufwerkbuchstabe*:\dmb\db2tx\dict
! oder
! *DB2TX_INSTOWNER* /db2tx/dicts

Die Dateien, die zu diesem Thesaurus gehören, werden `desthes.th1`, `desthes.th2`, ... und `desthes.th6`. genannt.

Um einen Thesaurus zu erstellen, definieren Sie zunächst seinen Inhalt in einer Datei. Es wird empfohlen, daß Sie ein neues Verzeichnis für jeden Thesaurus verwenden, den Sie definieren. Die Datei kann eine beliebige Erweiterung haben, mit Ausnahme von `th1` bis `th6`, die für das Thesaurus-Wörterverzeichnis verwendet werden. Wenn Sie dasselbe Verzeichnis für einen Ngram-Thesaurus verwenden, finden Sie in „Einen Ngram-Thesaurus erstellen“ auf Seite 41 weitere ausgeschlossene Dateierweiterungen.

Kompilieren Sie anschließend die Datei, indem Sie folgenden Befehl ausführen:

```
txthesc -f dateiname -c ccsid
```

! Momentan wird nur die CCSID 850 unterstützt.

Thesaurus-Konzepte

`txthesc` erstellt Thesaurus-Dateien mit dem Namen *dateiname* ohne Erweiterung sowie die Erweiterung `th1` bis `th6` in demselben Verzeichnis, in dem sich die Definitionsdatei befindet. Wenn dort bereits ein Thesaurus mit demselben Namen existiert, wird er ohne Warnung überschrieben.

In Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, finden Sie Informationen, wie ein Thesaurus in einer Abfrage verwendet wird.

Geben Sie den Inhalt eines Thesaurus unter Verwendung der Standard Generalized Markup Language (SGML) an. Das folgende Diagramm zeigt die Syntaxregeln, die beim Erstellen eines Thesaurus befolgt werden müssen.

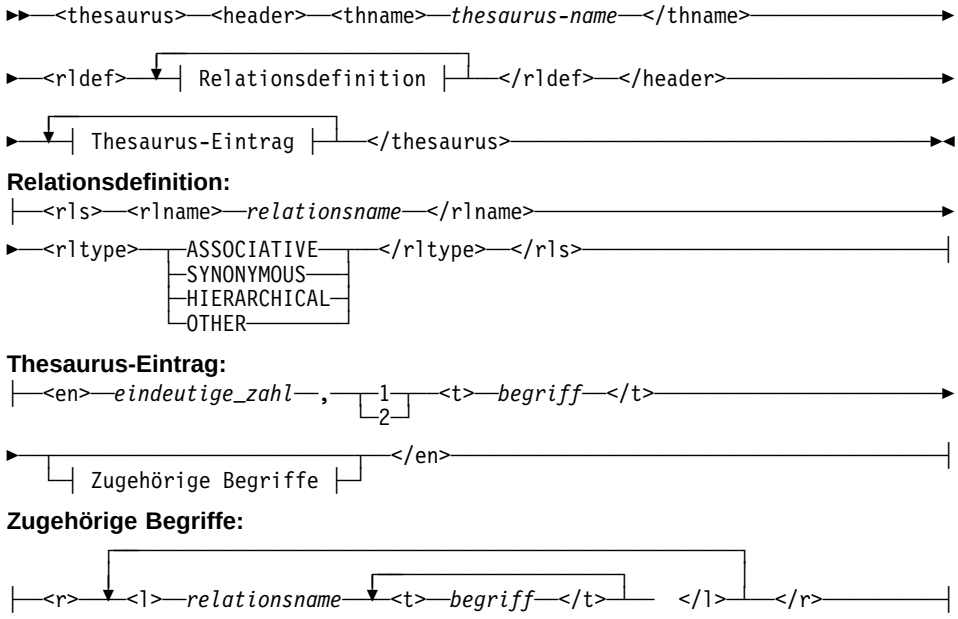


Abb. 6 auf Seite 39 zeigt die SGML-Definition des Thesaurus, der in Abb. 5 auf Seite 33 gezeigt wird.

```
<thesaurus>
<header>
<thname>thesc example thesaurus</thname>
<rldef>

<rls>
<rlname>Related Term</rlname>
<rltype>associative</rltype>
</rls>

<rls>
<rlname>Narrower Term</rlname>
<rltype>hierarchical</rltype>
</rls>

<rls>
<rlname>Instance</rlname>
<rltype>hierarchical</rltype>
</rls>

<rls>
<rlname>Synonym</rlname>
<rltype>synonymous</rltype>
</rls>
</rldef>
</header>

<en> 2, 1
<t>database management system</t>
<r>
  <l>Narrower Term
  <t>oo database management system</t>
  <t>relational database management system</t>
  </l>

  <l>Synonym
  <t>DBMS</t>
  </l>

  <l>Related Term
<t>document management system</t>
  </l>

  <l>Instance
<t>database</t>
  </l>
</r>
</en>
```

Abbildung 6 (Teil 1 von 2). Die Definition eines einfachen Thesaurus

```
<en> 5, 1
<t> relational database management system </t>
<r>
  <l>Narrower Term
  <t>object relational database management system</t>
  </l>
</r>
</en>

<en> 3, 1
  <t>object relational database management system</t>
<r>
  <l>Instance
  <t>DB2 Universal Database</t>
  </l>
</r>
</en>

<en> 6, 1
<t>object oriented database management system</t>
<r>
  <l>Narrower Term
  <t>object relational database management system</t>
  </l>
</r>
</en>

<en> 4, 1
<t>document management system</t>
<r>
  <l>Synonym
<t>library</t>
  </l>
</r>
</en>

<en> 9, 1
<t>library</t>
</en>

<en> 10, 1
<t>DB2 Unversal Database</t>
</en>

<en> 11, 1
<t>database</t>
</en>
</thesaurus>
```

Abbildung 6 (Teil 2 von 2). Die Definition eines einfachen Thesaurus

Einen Ngram-Thesaurus erstellen

Ein englisches Beispiel für eine Ngram-Thesaurus-Compiler-Eingabedatei, `desnthes.def`, ist im Verzeichnis für das Wörterverzeichnis des Installationspfades gespeichert. Das Verzeichnis für das Wörterverzeichnis auf OS/2- und Windows-Systemen ist

laufwerkbuchstabe`:\dmb\db2tx\dict`

Auf AIX-, HP-UX- und SUN-Solaris-Systemen ist das Verzeichnis für das Wörterverzeichnis

DB2TX_INSTOWNER `/db2tx/dicts`

Eine kompilierte Version dieses Beispiel-Thesaurus ist ebenfalls hier gespeichert. Die Dateien, die zu diesem Thesaurus gehören, werden `desnthes.<erweiterung>` genannt, mit der folgenden Erweiterung, wobei n eine Ziffer ist:

- Für Wörterverzeichnisdateien: wdf, wdv, grf, grv, MEY, ROS, NEY, SOS, lk*n*
- Für temporäre Dateien: : w*n*f, w*n*v, g*n*f, g*n*v, M!1, M!2, N!1, N!2, R!1, R!2, S!1, S!2, M*nn*, N*nn*, R*nn*, S*nn*, $00, $01, $10, $11, $20 und $21

Um einen Ngram-Thesaurus zu erstellen, definieren Sie zunächst seinen Inhalt in einer Definitionsdatei. Sie können mehrere Thesauri in demselben Verzeichnis speichern, aber es ist zu empfehlen, ein separates Verzeichnis für jeden Thesaurus einzurichten. Die Länge des Dateinamens ohne Erweiterung darf 8 Zeichen nicht überschreiten. Die Erweiterung ist wahlfrei, aber auf 3 Zeichen eingeschränkt und sollte sich von den im vorherigen aufgelisteten Erweiterungen unterscheiden.

Wenn Sie dasselbe Verzeichnis für andere Text Extender-Thesauri verwenden, verwenden Sie nicht die Erweiterungen, die unter „Einen Thesaurus erstellen“ auf Seite 37 aufgelistet sind.

Kompilieren Sie anschließend die Datei, indem Sie folgenden Befehl ausführen:

`txthesn -f` *definitionsdateiname* `-ccsid` *codepage*

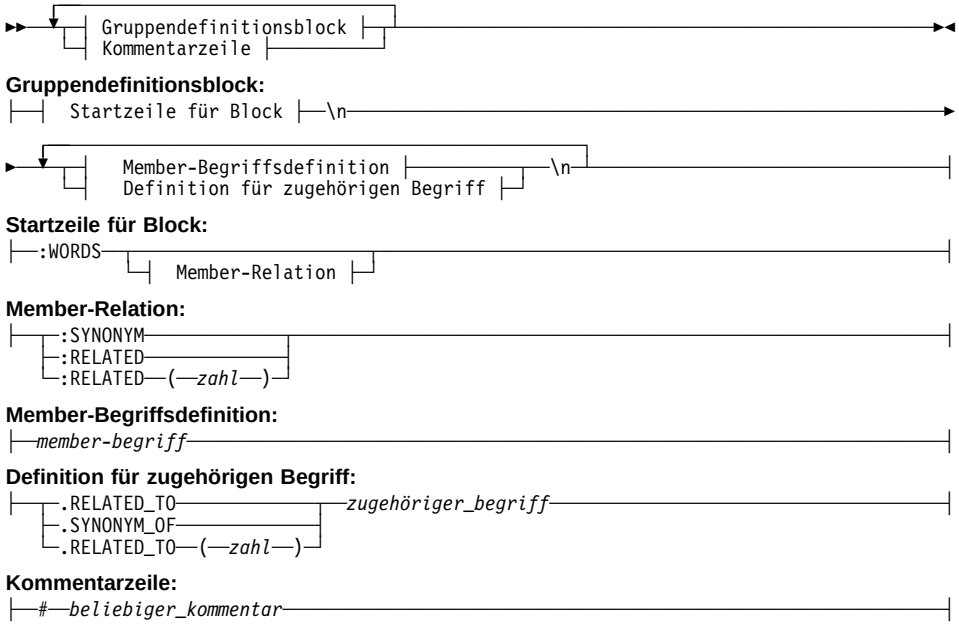
Die unterstützten Codepages werden in „Ngram-Index“ auf Seite 15 aufgelistet.

`txthesn` erstellt Thesaurus-Dateien mit dem Namen *definitionsdateiname* mit den im vorherigen aufgeführten Erweiterungen. Die Dateien werden in demselben Verzeichnis erstellt wie die Definitionsdatei. Wenn in diesem Verzeichnis bereits ein Thesaurus mit demselben Namen existiert, wird er ohne Warnung überschrieben.

Die unterstützten CCSIDs werden in der Beschreibung zu EhwCreateIndex im Buch *Text Search Engine: Programming Interfaces* aufgelistet.

Thesaurus-Konzepte

Geben Sie den Inhalt des Thesaurus unter Verwendung des folgenden Syntaxdiagramms an:



Jeder Member-Begriff muß jeweils in eine einzelne Zeile geschrieben werden. Vor jedem zugehörigen Begriff muß der Relationsname angegeben sein. Wenn die Member-Begriffe in Beziehung zueinander stehen, geben Sie eine Member-Relation an.

Die Länge der Member-Begriffe und der zugehörigen Begriffe ist auf 164 Zeichen eingeschränkt. Einzelbytezeichen und Doppelbytezeichen desselben Buchstabens werden als gleich angesehen. Groß- und Kleinbuchstaben werden nicht unterschieden. Ein Begriff kann ein Leerzeichen enthalten, aber es kann entweder das Einzelbytezeichen Punkt "." oder Doppelpunkt ":" verwendet werden.

Die benutzerdefinierten Relationen basieren alle auf der Relationsart *assoziativ*. Sie sind durch eindeutige Zahlen zwischen 1 und 128 gekennzeichnet.

Wenn eine Anwendung symbolische Namen für ihre Thesaurus-Relationen anstelle von Relationsname und Zahl verwenden soll, müssen Sie die Zuordnung selbst verwalten. Wenn beispielsweise die Relation OPPOSITE_OF als RELATED_TO(1) definiert wurde, muß die Anwendung diesen Namen dem internen Relationsnamen RELATED_TO(1) zuordnen. In Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, finden Sie Informationen, wie die Thesaurus-Erweiterung in einer Abfrage verwendet wird.

Kapitel 4. Verwaltung

Dieses Kapitel beginnt mit einem kurzen Abschnitt zur Verwaltung eines Text Extender-Servers, befaßt sich aber hauptsächlich mit der Verwaltung von Text Extender-Clients. Es beschreibt, wie Textdokumente für die Verwendung durch den Text Extender vorbereitet und wie Textindizes verwaltet werden. In diesem Kapitel wird davon ausgegangen, daß die Textdokumente, die Sie durchsuchen wollen, bereits in einer oder mehreren Tabellenspalten gespeichert sind.

Wenn Sie es in diesem Stadium vorziehen, die Suchfähigkeiten des Text Extender zu testen, indem Sie mit einer Beispieldatenbank arbeiten, die bereits vorbereitete Textdokumente enthält, können Sie zu „Eine Beispieldatenbank vorbereiten, die aktivierten Text enthält“ auf Seite 55 springen und anschließend mit Kapitel 5, „Mit Hilfe von Text Extender-UDFs suchen“ auf Seite 95, fortfahren.

! Der Beispiel-Wizard für die Indexierung und die Suche (nur NT)

! Für Windows NT-Benutzer wird der Indexierungs- und Such-Wizard des Text Extender
! als Beispiel dafür geliefert, wie Indizes für ausgewählte Tabellen oder Spalten erstellt
! und diese Indizes durchsucht werden. Dieser Beispiel-Wizard führt Sie durch die
! Schritte, die für die Suche nach ausgewählten Spalten einer Tabelle unter Verwendung
! der Suchfunktionen des Text Extender erforderlich sind.

! **Zum Starten des Beispiel-Wizard** geben Sie folgendes ein:

```
! txwizard
```

! Sie können den **Java-Quellencode** verwenden, der mit dem Beispiel-Wizard geliefert
! wird, um entweder Ihre eigene Verwaltungs- und Suchanwendung zu erstellen oder nur
! einen Eindruck zu erhalten, wie die Text Extender-Funktionen verwendet und innerhalb
! einer Java-Umgebung aufgerufen werden können. Der Java-Quellencode des Wizard
! ist in folgendem Verzeichnis gespeichert:

```
! %DMBMMPATH%\samples\wizard
```

Ein Text Extender-Exemplar erstellen

Bevor Sie die Arbeit mit dem Text Extender beginnen können, müssen Sie ein Text Extender-Exemplar erstellen. Jedes Exemplar, das Sie erstellen, bietet eine isolierte Verwaltungsumgebung, in der Sie Indizes verwalten und in separaten Verzeichnissen speichern können.

Um ein Exemplar in einer UNIX-Umgebung auf einem Einzelknoten zu erstellen, geben Sie folgenden Befehl ein:

```
txicrt  exemplarname db2_exemplarname
```

Dabei ist *exemplarname* die Benutzer-ID des Exemplareigners und *db2_exemplarname* der Name des DB2 UDB-Exemplars, das dem Text-Extender-Exemplar zugeordnet werden soll.

Einen Text Extender-Server verwalten

Unter Windows NT lautet der Befehl wie folgt:

```
txicrt db2_exemplarname
```

Zum Erstellen von Exemplaren in einer Umgebung mit mehreren Knoten siehe „TXICRT“ auf Seite 163.

Um ein Exemplar auf UNIX-Systemen zu löschen, geben Sie folgenden Befehl ein:

```
txidrop  exemplarname db2_exemplarname
```

! Um ein Exemplar auf Windows NT-Systemen zu löschen, geben Sie folgenden Befehl
! ein:

```
txidrop db2_exemplarname
```

! Um festzustellen, mit welchem Exemplar Sie arbeiten wollen, führen Sie folgende
! Schritte aus:

1. Setzen Sie die entsprechende Umgebungsvariable:

   ```
   set DB2INSTANCE = exemplarname
   ```

2. Starten Sie DB2 UDB.
3. Starten Sie den Text Extender.

Einen Text Extender-Server verwalten

Verwenden Sie die Verwaltungsbefehle des Text Extender-Servers, um folgende Schritte auszuführen:

- Starten des Servers
- Anzeigen des Status des Servers
- Stoppen des Servers

Den Server starten

Um den Server zu starten, melden Sie sich unter der Benutzer-ID des Exemplareigners (nur UNIX-Systeme) an und geben Sie den folgenden Befehl ein:

```
txstart
```

Den Status des Servers anzeigen

Um den Status des Servers anzuzeigen, geben Sie folgenden Befehl ein:

```
txstatus
```

Den Server stoppen

Um den Server zu stoppen, geben Sie folgenden Befehl ein:

```
txstop
```

Überblick über die Verwaltungs-Tasks auf dem Client

Der Text Extender muß bereits auf dem Server gestartet sein, bevor Sie die Verwaltungsbefehle des Clients verwenden können. Fordern Sie im Zweifelsfall einen Administrator auf, durch Eingabe des Befehls `txstatus` auf dem Server, auf dem das Text Extender-Exemplar installiert ist, zu prüfen, ob der Text Extender gestartet wurde.

Wenn Sie den Text Extender gerade installiert haben, führen Sie die Tasks in diesem Kapitel bis einschließlich „Eine Textspalte aktivieren“ auf Seite 60 in der angegebenen Reihenfolge aus. Der Rest der Abschnitte befaßt sich mit der Indexverwaltung.

! In diesem Kapitel benötigen Sie eine DB2 UDB-Datenbank, mit der Sie arbeiten. Wenn
! Sie noch keine Datenbank haben, erstellen Sie eine. Stellen Sie sicher, daß die
! Datenbank-CCSID eine der vom Text Extender unterstützten IDs ist, die in „CCSIDs“
! auf Seite 259 aufgelistet sind.

Zu jeder Task gehört eine Zusammenfassung, wann die Task auszuführen ist, welcher Befehl zu verwenden ist und welche Berechtigung erforderlich ist. Eine Beschreibung der Befehlsparameter finden Sie in Kapitel 7, „Verwaltungsbefehle für den Client“ auf Seite 127.

Sie führen die Verwaltungs-Tasks auf dem Client aus, indem Sie Befehle an der Eingabeaufforderung des Betriebssystems eingeben. Die Befehle ähneln DB2-Befehlen, aber statt ihnen `db2` voranzustellen, müssen Sie `db2tx` voranstellen.

Verwaltungsüberblick

- Die Verwaltung starten:

  Führen Sie diese Tasks vor allen anderen Verwaltungs-Tasks aus:

  1. Den Text Extender-Befehlszeilenprozessor starten (wahlfrei)
  2. Eine Verbindung zu einer Datenbank herstellen (wahlfrei)

- Textdokumente für die Suche vorbereiten:

  Führen Sie diese Tasks in der angegebenen Reihenfolge aus, bevor Sie eine Suche starten:

  1. Die Textkonfiguration ändern (wahlfrei)
  2. Die Stoppwort- und Abkürzungsdateien ändern (wahlfrei)
  3. ! Die Dokumentmodelldatei ändern (wahlfrei)
  4. Eine Beispieltabelle erstellen (wahlfrei)
  5. Eine Datenbank für die Verwendung durch den Text Extender aktivieren
  6. Eine Texttabelle für die Verwendung durch den Text Extender aktivieren (wahlfrei)

Überblick über die Verwaltungs-Tasks auf dem Client

7. Eine Textspalte für die Verwendung durch den Text Extender aktivieren
8. Externe Textdateien für die Verwendung durch den Text Extender aktivieren (wahlfrei)

- Den Textvorbereitungsprozeß umkehren:

  Eine Textspalte für die Verwendung durch den Text Extender inaktivieren

  Externe Textdateien für die Verwendung durch den Text Extender inaktivieren

  Eine Texttabelle für die Verwendung durch den Text Extender inaktivieren

  Eine Datenbank für die Verwendung durch den Text Extender inaktivieren
- Textindizes verwalten:

  Einen Index aktualisieren

  Die Einstellungen eines Indexes ändern

  Den Status eines Indexes zurücksetzen

  Indexereignisse löschen
- Informationen abrufen:

  Informationen zum aktivierten Status anzeigen

  Die Einstellungen der Umgebungsvariablen anzeigen

  Die Textkonfigurationseinstellungen anzeigen

  Den Indexstatus anzeigen

  Fehlerereignisse anzeigen

  Die Indexeinstellungen anzeigen

  Die Texteinstellungen für eine Spalte anzeigen
- Mit der Text Extender-Katalogsicht arbeiten
- Fehler-Trace ausführen
- Indizes und aktivierte Datenbanken sichern und zurückschreiben

Einführung

Hinweis für

UNIX-Benutzer

Wenn Sie kein Text Extender-Exemplareigner sind, finden Sie in Kapitel 14, „Konfiguration“ auf Seite 253, Informationen, wie Sie Ihr Profil einrichten.

Während dieser Verwaltungsschritte fällen Sie Entscheidungen auf der Basis Ihrer Text Extender-Konzeptkenntnisse. Bevor Sie beginnen, sollten mit folgenden Themen vertraut sein:

- Das Konzept von *Stoppwort*- und Abkürzungslisten und die Frage, ob Sie sie vor Beginn der Indexierung ändern wollen (siehe „Warum Textdokumente indexiert sein müssen“ auf Seite 11).
- Ob ein Index für die gesamte Texttabelle oder ein separater Index für jede Textspalte erstellt werden soll (siehe „Einen oder mehrere Textindizes für eine Tabelle erstellen“ auf Seite 16).
- Die CCSIDs, die Sprachen und die Formate der Texte, in denen Sie suchen wollen (siehe „Informationen zu Textdokumenten“ auf Seite 256), und die Textkonfigurationseinstellungen für CCSID, LANGUAGE und FORMAT, die die Standardwerte für diese Parameter setzen (siehe „Textmerkmale“ auf Seite 254).
- Der Textindextyp, den Sie verwenden (siehe Kapitel 2, „Einen Textindex planen“ auf Seite 11), und der Standardindextyp, der in den Textkonfigurationseinstellungen gesetzt ist (siehe „Indexmerkmale“ auf Seite 255).
- Das Verzeichnis, in dem Sie Indizes speichern wollen, und der Wert der Textkonfigurationseinstellung für DIRECTORY, der den Standardwert für diesen Parameter setzt (siehe „Indexmerkmale“ auf Seite 255).
- Der Standarddatenbankname in der DB2-Umgebungsvariablen DB2DBDFT (siehe „Umgebungsvariablen“ auf Seite 253).

Die Verwaltung starten

In diesem Abschnitt wird beschrieben, welche Schritte Sie bei Beginn jeder Verwaltungssitzung ausführen müssen.

Den Text Extender-Befehlszeilenprozessor starten (wahlfrei)

Zusammenfassung

Zeitpunkt	Bei Beginn jeder Verwaltungssitzung.
Befehl	db2tx
Berechtigung	Beliebig

Geben Sie den folgenden Befehl ein, um den Text Extender-Befehlszeilenprozessor zu starten:

```
db2tx
```

Die `db2tx`-Eingabeaufforderung wird angezeigt; alle nachfolgenden Befehle werden als Text Extender-Befehle interpretiert.

```
db2tx =>
```

Um diesen Modus zu verlassen, geben Sie QUIT ein.

Unter Windows 95 und Windows 98 können Sie den Befehl `db2tx` wie beschreiben verwenden, oder Sie können auf den Knopf 'Start' klicken, auf 'Programme' zeigen, auf 'DB2 Client für Windows 95/98' zeigen und anschließend auf 'DB2-Befehlsfenster' klicken. Wenn Sie diesen Schritt auslassen, können Sie Text Extender-Befehle direkt von der Eingabeaufforderung des Betriebssystems aus eingeben, indem Sie das Präfix `db2tx` verwenden. Hier ist ein Beispiel für einen Befehl, der von der Eingabeaufforderung des Betriebssystems aus eingegeben wurde:

```
db2tx enable database
```

Hinweis

Wenn Sie die Verwaltung von einem Client aus ausführen, müssen Sie in der Text Extender-Befehlsumgebung sein und das vollständige Format der Anweisung CONNECT verwenden, indem Sie Ihre Benutzer-ID und Ihr Kennwort angeben.

! `db2tx => connect to` *`datenbankname`* `user` *`benutzer-ID`* `using` *`kennwort`*

Eine Verbindung zu einer Datenbank herstellen

Zusammenfassung

Zeitpunkt Wahlfrei. Zum Starten eine Verwaltungssitzung mit einer anderen Datenbank als der in der Umgebungsvariablen DB2DBDFT angegebenen Standarddatenbank.

Befehl CONNECT

Berechtigung CONNECT für die Datenbank.

Bevor Sie weitere Verwaltungsbefehle in der Text Extender-Sitzung ausgeben können, müssen Sie eine Verbindung zu einer Datenbank herstellen. Sie können explizit eine Verbindung zu einer Datenbank herstellen, indem Sie den Text Extender-Befehl CONNECT TO verwenden. Wenn Sie einen Verwaltungsbefehl ausgeben, ohne mit einer Datenbank verbunden zu sein, stellt der Text Extender eine Verbindung zur Standarddatenbank her, die in der Umgebungsvariablen DB2DBDFT angegeben ist.

```
db2tx "CONNECT TO beispieldatenbank"
```

Die folgenden Informationen werden angezeigt (hier die AIX-Version):

```
>------------Datenbankverbindungsinformationen-----------<
Databankprodukt      = DB2/6000 05.02.0000
Benutzer             = benutzername
Aliasname für DB     = beispieldatenbank
```

In den späteren Verwaltungsbeispielen wird davon ausgegangen, daß Sie mit der Datenbank verbunden sind, die die Beispieltabelle enthält, die in „Eine Beispieltabelle erstellen“ auf Seite 54 beschrieben ist.

Anmerkung: Wenn Sie anschließend einen Text Extender-Befehl außerhalb des Befehlsprozessors eingeben, indem Sie das Präfix `db2tx` verwenden, wird die aktuelle Datenbankverbindung unterbrochen und eine neue Verbindung zur Standarddatenbank wird hergestellt.

Wenn Sie einen OS/2-Client verwenden und zum ersten Mal eine Verbindung zu einer Datenbank herstellen, nachdem Sie das System gestartet haben, werden Sie von der Benutzerprofilverwaltung (UPM) aufgefordert, eine Benutzer-ID und ein Kennwort einzugeben. Wenn Sie bereits angemeldet sind, übernimmt die UPM die Benutzer-ID und das Kennwort aus dieser Sitzung und versucht, eine Verbindung zur fernen Datenbank auf dem Server herzustellen.

Textdokumente für die Suche vorbereiten

Textdokumente für die Suche vorbereiten

In diesem Abschnitt wird beschrieben, wie eine Datenbank vorbereitet wird, so daß ihre Texttabellen vom Text Extender durchsucht werden können. Zu diesen Schritten gehören:

1. Die Textkonfiguration ändern (wahlfrei)
2. Die Stoppwort- und Abkürzungsdateien ändern (wahlfrei)
3. Die Dokumentmodelldatei ändern (wahlfrei)
4. Eine Beispieltabelle erstellen (wahlfrei)
5. Eine Datenbank für die Verwendung durch den Text Extender aktivieren
6. Eine Texttabelle für die Verwendung durch den Text Extender aktivieren (wahlfrei)
7. Eine Textspalte für die Verwendung durch den Text Extender aktivieren
8. Den Text Extender für die Suche in externen Textdateien aktivieren (wahlfrei)

Ein Auszug aus der Beispieltabelle wird in Tabelle 6 auf Seite 96 gezeigt.

Wenn Sie den Text Extender gerade installiert haben, führen Sie diese Schritte in der angegebenen Reihenfolge aus.

Die Textkonfiguration ändern

Zusammenfassung

Zeitpunkt Wenn Sie unterschiedliche Standardeinstellungen festlegen wollen, die beim Erstellen und Aktualisieren eines Indexes verwendet werden.

Befehl CHANGE TEXT CONFIGURATION

Berechtigung SELECT

Textdokumente für die Suche vorbereiten

Wenn Sie einen Index erstellen, werden die folgenden Parameter (die in „Textkonfigurationseinstellungen“ auf Seite 254 beschrieben sind) für diesen Index gesetzt:

- CCSID (ID für codierten Zeichensatz)
- Sprache
- Format
- Indextyp
- Aktualisierungsfrequenz
- Indexverzeichnis
- Option UPDATE INDEX
- COMMIT-Anzahl

Wenn der Text Extender das erste Mal installiert wird, werden die Standardwerte für diese Einstellungen in der *Textkonfiguration* festgelegt. Um die aktuellen Werte der Textkonfiguration aufzurufen, siehe „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 84.

Um die Textkonfiguration, die beim Erstellen von Indizes als Standardwert verwendet wird, zu ändern, geben Sie folgendes ein:

```
db2tx "CHANGE TEXT CFG USING einstellungen"
```

Beispiele

Um den Standardindextyp und das Standardindexverzeichnis für zukünftige Indizes zu ändern, geben Sie folgendes ein:

```
db2tx "CHANGE TEXT CONFIGURATION USING
                  INDEXTYPE   precise
                  INDEXOPTION normalized
                  DIRECTORY   DB2TX_INSTOWNER/db2tx/indexes"
```

Um die Standardaktualisierungsfrequenz für Indizes zu ändern, so daß sie um 12:00 oder 15:00 Uhr von Montag bis Freitag aktualisiert werden, wenn mindestens 100 Textdokumente in der Warteschlange stehen, geben Sie folgendes ein:

```
db2tx "CHANGE TEXT CONFIGURATION USING
                  UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)"
```

Um das regelmäßige Aktualisieren eines Indexes zu stoppen, geben Sie folgendes ein:

```
db2tx "CHANGE TEXT CONFIGURATION USING
                  UPDATEFREQ none"
```

Textdokumente für die Suche vorbereiten

Die Stoppwort- und Abkürzungsdateien ändern

Zusammenfassung

Zeitpunkt	Wenn möglich, nur einmal, wenn der Text Extender das erste Mal installiert wird. Wahlfrei.
Befehl	Ihr eigener Editorbefehl
Berechtigung	Keine

Pro Sprache gibt es eine Stoppwortdatei und eine Abkürzungsdatei. Um die Folgen beim Editieren dieser Dateien zu verstehen, lesen Sie „Warum Textdokumente indexiert sein müssen“ auf Seite 11.

Hinweis

Bevor Sie mit dem Editieren einer dieser Dateien beginnen, erstellen Sie eine Sicherungskopie.

Die Stoppwort- und Abkürzungsdateien befinden sich auf OS/2- und Windows-Systemen in folgendem Verzeichnis:

laufwerkbuchstabe:\dmb\db2tx\dict

Auf AIX-, HP-UX- und SUN-Solaris-Systemen befinden sich die Dateien in folgendem Verzeichnis:

DB2TX_INSTOWNER/db2tx/dicts

Verwenden Sie Ihren Editor, um diese Dateien zu bearbeiten. Die Dateien verwenden die CCSID 850; stellen Sie also sicher, daß Ihre Anwendungs-CCSID ebenfalls auf 850 gesetzt ist, bevor Sie beginnen.

Löschen Sie Wörter und Abkürzungen, die indexiert werden sollen. Fügen Sie Wörter hinzu, die nicht indexiert werden sollen.

Die Dokumentmodelldatei ändern

Zusammenfassung

Zeitpunkt	Wenn Sie die Vorteile der Struktur eines Dokuments nutzen wollen. Wahlfrei.
Befehl	Ihr eigener Editorbefehl
Berechtigung	Keine

Sie können eine Suche auf bestimmte Abschnitte von Dokumenten beschränken. Dieses Konzept sowie die Verwendung einer Dokumentmodelldatei zum Aktivieren des Text Extender, damit er solche Abschnitte erkennt, werden in „Die Vorteile der Struktur eines Dokument nutzen“ auf Seite 68 beschrieben.

Wenn ein Server-Exemplar erstellt wird, wird ein Beispiel einer Dokumentmodelldatei, *desmodel.ini*, im Verzeichnis des Server-Exemplars erstellt. In derselben Datei befindet sich auch ein Beispiel eines Dokumentmodells für HTML-Dokumente.

Verwenden Sie Ihren Editor, um Dokumentmodelldateien zu editieren.

Später, wenn Sie die Textspalte, die die Dokumente enthält, aktivieren, müssen Sie `INDEXPROPERTY SECTIONS_ENABLED` angeben.

Textdokumente für die Suche vorbereiten

Eine Beispieltabelle erstellen

Zusammenfassung

Zeitpunkt Wahlfrei. Zum Erstellen einer Tabelle, die in späteren Beispielen verwendet werden kann.

Befehl db2 "CREATE TABLE ..."

Berechtigung SYSADM oder DBADM

In diesem Abschnitt wird beschrieben, wie die Tabelle DB2TX.SAMPLE erstellt wird, die in vielen Beispielen in diesem Kapitel und in Kapitel 5, „Mit Hilfe von Text Extender-UDFs suchen“ auf Seite 95, verwendet wird.

Als Alternative können Sie TXVERIFY und TXSAMPLE ausführen, um automatisch eine Beispieltabelle zu erstellen, die für jeden Indextyp aktiviert werden kann. Siehe „Eine Beispieldatenbank vorbereiten, die aktivierten Text enthält“ auf Seite 55. In diesem Fall können Sie jedoch nicht länger den Tabellennamen DB2TX.SAMPLE in den Beispielen der Verwaltungsbefehle zum Aktivieren einer Texttabelle und zum Aktivieren einer Textspalte verwenden.

Wenn Sie sich entscheiden, nicht TXSAMPLE zum automatischen Erstellen der Beispieltabelle zu verwenden, erstellen Sie eine Befehlsprozedur, die folgende Schritte ausführt, und führen Sie die Befehlsprozedur aus:

1. Zum Erstellen der Beispieltabelle:

```
db2 "CREATE TABLE db2tx.sample (
        docid      VARCHAR(30),
        author     VARCHAR(50),
        subject    VARCHAR(100),
        date       TIMESTAMP,
        comment    LONG VARCHAR)"
```

2. Zum Füllen der Tabelle mit Daten:

```
db2 "import from \"dessamp.imp\"
of del modified by DELPRIORITYCHAR
insert into db2tx.sample"
```

   Die Datei DESSAMP.IMP befindet sich im Unterverzeichnis `/samples` des Verzeichnisses, in dem der Text Extender installiert ist.

Eine Beispieldatenbank vorbereiten, die aktivierten Text enthält

Der Text Extender bietet Dienstprogramme zum Vorbereiten einer Beispieldatenbank. Sie sind nützlich für das schnelle Vorbereiten von Text zum Testen der Suchfähigkeiten des Text Extender sofort nach der Installation und zu allgemeinen Testzwecken zu einem beliebigen späteren Zeitpunkt.

Diese Dienstprogramme befinden sich in den folgenden Verzeichnissen:

- Unter OS/2 und Windows NT: *DMBMMPATH*`\samples`

  wobei *DMBMMPATH* der Installationspfad ist.
- Unter AIX, HP-UX und SUN-Solaris: `$HOME/db2tx/sample`

Um eine Beispieldatenbank vorzubereiten, führen Sie entweder auf dem Server oder auf einer Client-Workstation folgende Schritte aus:

1. Führen Sie an der Eingabeaufforderung des Betriebssystems folgenden Befehl aus:

   ```
   TXVERIFY datenbankname  [benutzer-ID] [kennwort]
   ```

   Dieser Befehl erstellt und aktiviert eine Beispieldatenbank.
2. Führen Sie an der Eingabeaufforderung des Betriebssystems folgenden Befehl aus:

   ```
   TXSAMPLE datenbankname [benutzer-ID] [kennwort]
   ```

   Die Benutzer-ID und das Kennwort sind nur erforderlich, wenn Sie an einer Client-Workstation arbeiten.

   Dieser Befehl führt folgende Schritte aus:

   a. Er stellt eine Verbindung zur angegebenen Datenbank her.
   b. Er erstellt die folgende Tabelle:
      - `db2tx.sample_p`

   c. Er importiert englische Beispieldokumente zum Auffüllen der Tabelle.
   d. Er aktiviert die Textspalte in der Tabelle mit den folgenden Indextypen:
      präzise (PRECISE)
      präzise normalisert (PRECISE NORMALIZED)
      linguistisch (LINGUISTIC)
      linguistisch abschnittsfähig (LINGUISTIC SECTIONS_ENABLED)
      Ngram (NGRAM)
      Ngram unter Beachtung der Groß-/Kleinschreibung (NGRAM CASE_ENABLED)
      dual (DUAL)
   e. Er wartet, daß der Textindex erstellt wird.

Textdokumente für die Suche vorbereiten

Eine Datenbank aktivieren

Zusammenfassung

Zeitpunkt Einmal für jede Datenbank, die Textspalten enthält, die durchsucht werden sollen.

Befehl ENABLE DATABASE

Berechtigung SYSADM oder DBADM

Um die verbundene Datenbank zu aktivieren, geben Sie folgendes ein:

```
db2tx "ENABLE DATABASE"
```

Dieser Befehl hat keine Parameter. Er bereitet eine Datenbank für die Verwendung durch den Text Extender vor.

Dieser Befehl deklariert außerdem benutzerdefinierte Funktionen (UDFs) und benutzerdefinierte eindeutige Typen (UDTs) für DB2. Hierbei handelt es sich um SQL-Funktionen, die Sie später für die Suche nach Text verwenden. Sie werden in Kapitel 9, „UDTs und UDFs“ auf Seite 179, beschrieben. Diese Deklarationen gelten für alle zukünftigen Sitzungen.

Eine Katalogsicht, TEXTINDEXES, wird erstellt, die die aktivierten Textspalten protokolliert. Siehe „Mit der Text Extender-Katalogsicht arbeiten“ auf Seite 89.

Dieser Befehl erstellt Textkonfigurationsdaten für die Datenbank. Dazu gehören Standardwerte für den Index, den Text und Verarbeitungsmerkmale. Sie werden in „Textkonfigurationseinstellungen“ auf Seite 254 beschrieben.

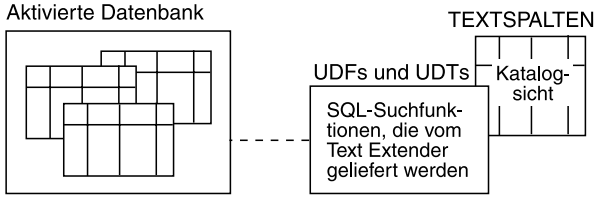


Abbildung 7. Eine Datenbank aktivieren

Wenn eine Datenbank aktiviert wurde, bleibt sie aktiviert, bis Sie sie inaktivieren. Um die Änderungen, die durch ENABLE DATABASE hergerufen wurden, umzukehren, gehen Sie zu „Eine Datenbank inaktivieren“ auf Seite 76.

Hinweise

1. Wenn die Umgebungsvariable DB2TX_INSTOWNER verwendet wird, muß sie auf den Namen des Exemplareigners gesetzt werden, bevor die Datenbank aktiviert wird. Dies ist besonders wichtig für UNIX-Benutzer, da unter UNIX diese Variable standardmäßig gesetzt wird.
2. Wenn Sie sich später entscheiden, eine aktivierte Datenbank zu löschen, sollten Sie sie zunächst inaktivieren, um sicherzustellen, daß die deklarierten UDFs, die Katalogsicht usw. gelöscht sind.

Eine Texttabelle aktivieren

Zusammenfassung

Zeitpunkt	Wahlfrei. Einmal, um einen gemeinsamen Index für alle Textspalten in der Tabelle zu erstellen.
Befehl	ENABLE TEXT TABLE
Berechtigung	ALTER, SELECT, UPDATE für die Tabelle

Dieser Schritt legt fest, ob Sie einen gemeinsamen Index für alle Textspalten in der Tabelle haben oder mehrere Indizes, d.&rbl;h. einen separaten Index für jede Textspalte. Weitere Informationen finden Sie in „Einen oder mehrere Textindizes für eine Tabelle erstellen“ auf Seite 16. Um einen gemeinsamen Index zu erstellen, führen Sie ENABLE TEXT TABLE aus und führen Sie anschließend ENABLE TEXT COLUMN für jede Textspalte aus. Um separate Indizes zu erstellen, überspringen Sie ENABLE TEXT TABLE und führen Sie nur ENABLE TEXT COLUMN für jede Textspalte aus. Dies wird in Abb. 8 und Abb. 9 auf Seite 60 gezeigt.

Während dieses Schritts erstellt der Text Extender einen leeren Textindex, der von allen nachfolgend aktivierten Textspalten gemeinsam benutzt wird. Sie geben den Indextyp, wie oft der Index aktualisiert werden soll und in welchem Verzeichnis der Index gespeichert werden soll, an. Standardwerte für alle Parameter, die Sie nicht angeben, werden aus den Textkonfigurationseinstellungen übernommen.

Bei den Beispielen in den anderen Kapiteln wird davon ausgegangen, daß der Indextyp DUAL ist.

Textdokumente für die Suche vorbereiten

Hinweis

Wenn eine Einstellung, z.&rbl;B. die Indexaktualisierungsfrequenz, für die meisten Texttabellen gleich sein soll, ist es möglicherweise besser, Textkonfigurationsdaten zu verwenden, um Standardeinstellungen anzugeben. Siehe „CHANGE TEXT CONFIGURATION“ auf Seite 130.

Dieser Schritt erstellt außerdem eine leere Protokolltabelle zum Aufzeichnen, welche Dokumente in der Tabelle hinzugefügt, geändert oder gelöscht werden. Auslöser werden erstellt, um die Protokolltabelle auf dem aktuellsten Stand zu halten.

Sie können ENABLE TEXT TABLE nicht für eine Tabelle ausführen, die bereits eine Textspalte enthält, die für den Text Extender aktiviert wurde.

Um einen Index, der mit ENABLE TEXT TABLE erstellt wurde, zu löschen, siehe „Eine Texttabelle inaktivieren“ auf Seite 76.

Hinweis

Wenn Sie sich später entscheiden, eine aktivierte Texttabelle zu löschen, sollten Sie sie zunächst inaktivieren, um sicherzustellen, daß der Index, die Protokolltabelle usw. gelöscht sind.

Beispiele

Das folgende Beispiel aktiviert die Texttabelle `DB2TX.SAMPLE`:

```
db2tx "ENABLE TEXT TABLE db2tx.sample"
```

Standardwerte für die Indexmerkmale werden aus den Textkonfigurationseinstellungen übernommen.

Im nächsten Beispiel werden die Merkmale des allgemeinen Indexes, der für die Tabelle erstellt wird, explizit gesetzt.

Für einen UNIX-Server:

```
db2tx "ENABLE TEXT TABLE    db2tx.sample
                INDEXTYPE  linguistic
                UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)
                DIRECTORY  DB2TX_INSTOWNER/db2tx/indexes"
```

Für einen OS/2- oder Windows NT-Server:

```
db2tx "ENABLE TEXT TABLE    db2tx.sample
                INDEXTYPE  linguistic
                UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)
                DIRECTORY  \dmb\db2tx\indexes"
```

Das Beispiel setzt den Index und das Indexverzeichnis und setzt anschließend die Indexaktualisierungsfrequenz so, daß der Index um 12:00 oder 15:00 Uhr von Montag bis Freitag aktualisiert wird, wenn mindestens 100 Textdokumente in der Warteschlange stehen.

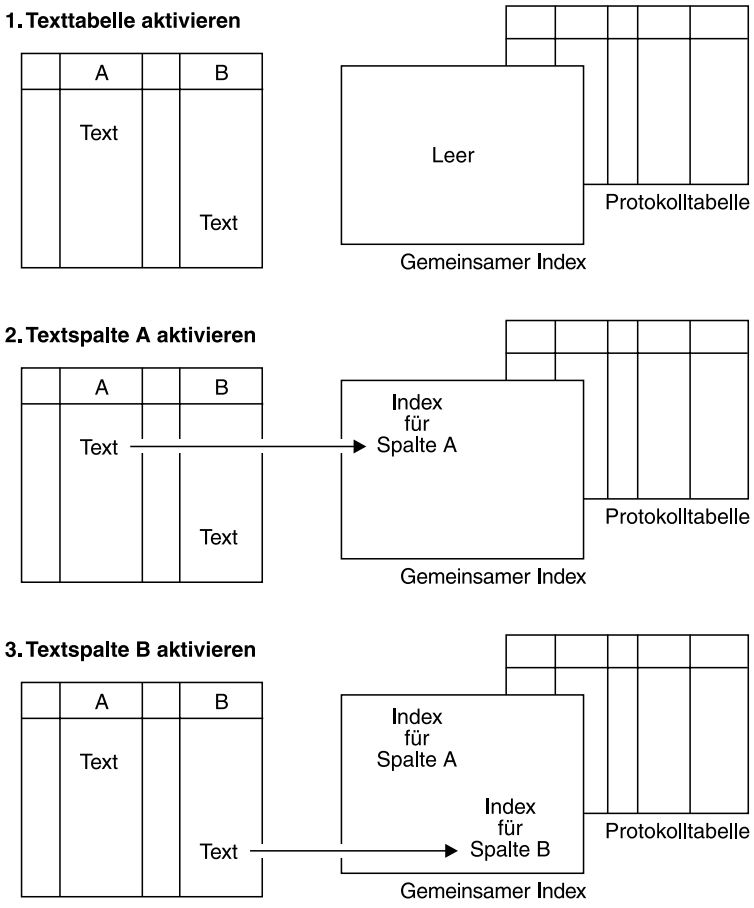


Abbildung 8. Einen gemeinsamen Index für alle Textspalten in einer Tabelle erstellen

Textdokumente für die Suche vorbereiten

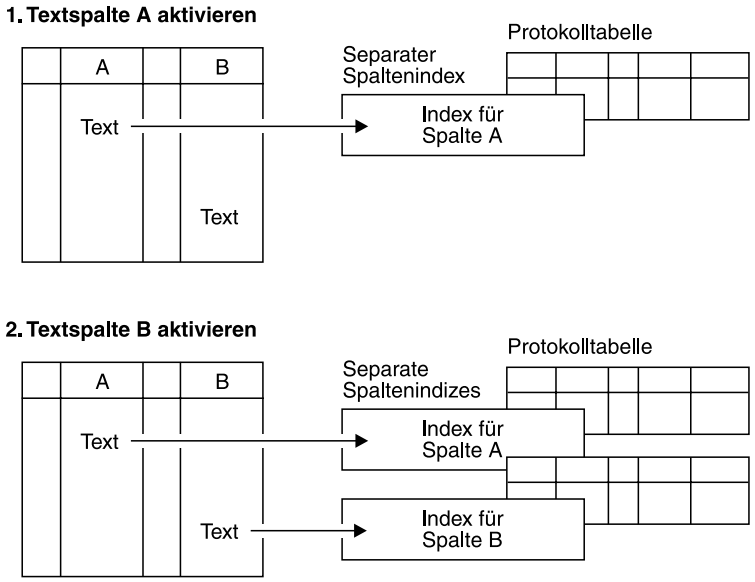


Abbildung 9. Einen separaten Index für jede Textspalte erstellen

Eine Textspalte aktivieren

Zusammenfassung

Zeitpunkt Einmal für jede Spalte, die zu durchsuchenden Text enthält.

Befehl ENABLE TEXT COLUMN

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Hinweis

Wenn eine Einstellung, z.&rbl;B. die Indexaktualisierungsfrequenz, für die meisten Textspalten gleich sein soll, verwenden Sie die Textkonfigurationsdaten, um Standardeinstellungen anzugeben.

Um die Änderungen, die durch ENABLE TEXT COLUMN hervorgerufen wurden, umzukehren, verwenden Sie den Befehl DISABLE TEXT COLUMN. Um alle aktivierten Textspalten in einer Tabelle zu inaktivieren, verwenden Sie den Befehl DISABLE TEXT TABLE.

! Wenn Sie eine Textspalte aktivieren, wird eine Kennungsspalte zur Tabelle hinzugefügt,
! werden Dokumentinformationen (Format, Sprache, CCSID) gesetzt, wird eine Protokoll-
! tabelle sowie ein Index erstellt.

Eine Kennungsspalte wird hinzugefügt

Während dieses Schritts fügt der Text Extender eine 60-Byte-VARCHAR-Kennungsspalte hinzu - eine Spalte, die Kennungen enthält, die der zu aktivierenden Textspalte zugeordnet sind. Kennungen enthalten Informationen zum Text in der zugeordneten Textspalte und in den zugeordneten externen Dateien. Zu diesen Informationen gehört eine eindeutige Dokument-ID, die Sprache, das Format und die CCSID des Dokuments sowie der Indexname. Sie werden in „Die Beispieltabelle DB2TX.SAMPLE“ auf Seite 96 beschrieben.

Hinweis

Diese Version des Text Extender kann nicht mit Indizes arbeiten, die unter Verwendung eines früheren Releases erstellt wurden. Sie müssen alle Text Extender-fähigen Datenbanken und Tabellen inaktivieren und anschließend erneut aktivieren.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
Daten	Daten	Daten	Daten	Text

Abbildung 10. Struktur der Tabelle DB2TX.SAMPLE - vor dem Aktivieren

Die Spalte mit Textblocken ist die Spalte COMMENT. Bevor Sie den Text in dieser Spalte durchsuchen können, müssen Sie die Datenbank und die Spalte COMMENT für die Verwendung durch den Text Extender vorbereiten.

Nach diesem Vorbereitungsschritt enthält die Tabelle DB2TX.SAMPLE eine zusätzliche Spalte für Kennungen.

DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT	COMMENTHANDLE
Daten	Daten	Daten	Daten	Text	Textkennungen

Abbildung 11. Struktur der Tabelle DB2TX.SAMPLE - nach dem Aktivieren

Anmerkung: Wenn Sie anschließend nach Text suchen, geben Sie als zu durchsuchende Spalte die Kennungsspalte, nicht die Textspalte an.

Textdokumente für die Suche vorbereiten

Die Dokumentinformationen werden gesetzt

Sie geben den Typ von Textdokumenten an, die Sie normalerweise in dieser Textspalte speichern: deren Format (z.&rbl;B. ASCII), deren Sprache und deren CCSID. Standardwerte für diese Informationen können in den Textkonfigurationseinstellungen angegeben werden. Siehe „Textkonfigurationseinstellungen“ auf Seite 254.

Eine Protokolltabelle wird erstellt

! Während dieses Schritts wird eine Protokolltabelle und eine Sicht mit dem Namen
! LOGIXnnnnnn erstellt, wobei *IXnnnnnn* der Indexname ist (aus der Katalogsicht ver-
! fügbar). Wenn in der Textkonfiguration ein Standardtabellenbereich angegeben ist, wird
! die Protokolltabelle dort gespeichert. Andernfalls wird sie im Standardtabellenbereich
! des DB2-Systems gespeichert. Um die Leistung und die Nutzung des Plattenspeicher-
! platzes zu optimieren, können Sie einen anderen Tabellenbereich angeben, der für die
! Protokolltabellen verwendet werden soll.

Auslöser werden ebenfalls erstellt, die immer Informationen zur Protokolltabelle hinzufügen, wenn ein Dokument in der Spalte hinzugefügt oder geändert wird. Diese Informationen führen dazu, daß beim nächsten Indexieren diese Dokumente indexiert werden.

Wenn externe Dateien hinzugefügt oder geändert werden, werden diese Änderungen von den Auslösern nicht erkannt. Um in solchen Fällen die Auslöser zu veranlassen, die Informationen zur Protokolltabelle hinzuzufügen, verwenden Sie eine Anweisung UPDATE, wie im Beispiel in „Einen Index für externe Dateien aktualisieren“ auf Seite 78 gezeigt.

Wenn während der Indexierung Fehler auftreten, d.&rbl;h., wenn z.&rbl;B. ein Dokument, das für die Indexierung in der Warteschlange steht, nicht gefunden werden kann, werden sogenannte *Fehlerereignisse* zur Protokolltabelle hinzugefügt und können, wie in „Fehlerereignisse anzeigen“ auf Seite 86 beschrieben, angezeigt werden.

Hinweis

Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 63.

In partitionierten Datenbanken ist jede Tabelle einem Tabellenbereich und einer Knotengruppe zugeordnet. Es ist wichtig, daß die Protokolltabelle einem Tabellenbereich zugeordnet ist, der zu derselben Knotengruppe gehört wie die aktivierte Benutzertabelle. Der Text Extender überprüft dies während der Ausführung des Befehls ENABLE.

Ein Index wird erstellt

Wenn Sie für jede Textspalte einen separaten Index haben wollen, d. h., wenn Sie den Schritt ENABLE TEXT TABLE übersprungen haben, erstellt der Text Extender einen separaten Index für jede Textspalte während dieses Schritts. Sie geben den Indextyp, wie oft der Index aktualisiert werden soll und in welchem Verzeichnis der Index gespeichert werden soll, an. Wenn Sie auf der anderen Seite einen Index für die gesamte Tabelle haben wollen, haben Sie bereits den Befehl ENABLE TEXT TABLE ausgeführt und die Indexparameter angegeben; sie werden ignoriert, wenn Sie sie an dieser Stelle wiederholen.

Verwenden Sie das Schlüsselwort UPDATEINDEX, um festzulegen, ob das Indexieren der Textdokumente in der angegebenen Textspalte sofort beginnen soll oder bei der nächsten geplanten regelmäßigen Indexierung. Wenn Sie dieses Schlüsselwort nicht verwenden, wird der Wert verwendet, der in den Textkonfigurationseinstellungen angegeben ist.

! **Erstellen von verschiedenen Typen für eine Textspalte** Sie können mehr als einen
! Index für eine Textspalte erstellen. Dies kann nützlich sein, wenn Sie beispielsweise die
! linguistische Suche und die Suche nach grober Übereinstimmung für dieselbe Text-
! spalte zulassen wollen, indem Sie die Textspalte unterschiedlichen Indextypen,
! z.&rbl;B. linguistische und Ngram-Indizes, zuordnen. Führen Sie dazu ENABLE TEXT
! COLUMN erneut aus und geben Sie nicht nur den zusätzlich zu erstellenden Indextyp
! an, sondern auch einen eindeutigen Namen für die Kennungsspalte.

Hinweis

Wenn Sie eine partitionierte Datenbank verwenden und Sie sich entschieden haben, die Konfiguration einer Knotengruppe zu ändern und eine Neuverteilung der Tabellenzeilen zu starten, müssen Sie die Textindizes und Protokolltabellen erneut erstellen, d.&rbl;h., Sie müssen alle Text Extender-fähigen Tabellen dieser Knotengruppe inaktivieren und erneut aktivieren.

Eine Textspalte in einer großen Tabelle aktivieren

Wenn Sie mit einer Tabelle arbeiten, die große Zeilenlängen hat, beachten Sie, daß durch das Aktivieren einer Textspalte eine Kennungsspalte vom Typ DB2TEXTH (VARCHAR 60) hinzugefügt wird. Ebenso wird durch das Aktivieren einer externen Datei eine Kennungsspalte vom Typ DB2TEXTFH (VARCHAR 210) hinzugefügt. Dies könnte von entscheidender Bedeutung sein, wenn die Tabelle ihre maximale Zeilenlänge erreicht, die von DB2 festgelegt ist.

Wenn Sie eine Textspalte in einer großen Tabelle aktivieren, verwenden Sie das DB2 UDB-Dienstprogramm REORGANIZE, um zu prüfen, ob die Tabelle reorganisiert werden muß.

Textdokumente für die Suche vorbereiten

Wenn Sie eine große Tabelle das erste Mal aktivieren, wird durch die folgenden Schritte die Indexierung beschleunigt:

1. Aktivieren Sie die Tabelle unter Verwendung der Option NOUPDATE. Hierdurch werden die Kennungen erstellt, aber die Dokumente noch nicht indexiert.
2. Reorganisieren Sie die Tabelle unter Verwendung des DB2 UDB-Dienstprogramms REORGANIZE.
3. Erstellen Sie den Index, indem Sie UPDATE INDEX ausführen.

Wenn Sie eine Textspalte oder externe Dateien aktivieren, fügt der Text Extender eine Kennungsspalte zur Tabelle hinzu und initialisiert die Kennungswerte, wodurch DB2 UDB-Protokolleinträge geschrieben werden. Wenn eine ungewöhnlich große Anzahl an Protokolleinträgen geschrieben wird, verfügt DB2 UDB möglicherweise nicht mehr über genügend Speicherbereich.

Es gibt zwei Möglichkeiten, diese Situation zu beheben; die erste ist aus Leistungsgründen zu bevorzugen:

- Erhöhen Sie den verfügbaren Protokollbereich, indem Sie den DB2 UDB-Befehl UPDATE DB CFG verwenden, um die Datenbankkonfigurationsparameter für LOGPRIMARY, LOGSECOND und LOGFILSIZ zu ändern. Die folgenden Werte können erfahrungsgemäß genommen werden; sie müssen Sie aber möglicherweise an Ihre Installation anpassen.

  ```
  LOGSECOND     50
  ```

  Stellen Sie sicher, daß die Summe aus LOGPRIMARY und LOGSECOND nicht größer als 128 ist. Sie sollten außerdem die Größe des Anwendungsfreispeichers erhöhen.

  ```
  APPLHEAPSZ   512
  ```

- Erzwingen Sie, daß DB2 UDB einen Zwischenbefehl COMMIT ausführt, indem Sie den Konfigurationsparameter COMMITCOUNT verwenden, der in Kapitel 14, „Konfiguration“ auf Seite 253, beschrieben wird. Der von Ihnen angegebene Wert gibt an, nach wie vielen Einfügungs- oder Aktualisierungsanweisungen der Text Extender eine DB2 UDB-Anweisung COMMIT ausgibt. Hierdurch wird die erforderliche Größe für Protokolltabellen verringert, allerdings wird auch die Zeit erhöht, die für den Aktivierungsschritt erforderlich ist.

Beispiele

Das folgende Beispiel aktiviert die Textspalte COMMENT in der Tabelle DB2TX.SAMPLE und ordnet der Kennungsspalte, die erstellt wird, den Namen COMMENTHANDLE zu:

```
db2tx "ENABLE TEXT COLUMN    db2tx.sample    comment
                   HANDLE    commenthandle"
```

Standardwerte für die Text- und Indexmerkmale werden aus den Textkonfigurationseinstellungen übernommen.

Im nächsten Beispiel werden die Werte für den Typ von Dokumenten, die in der Spalte COMMENT sind, explizit gesetzt. Standardwerte für die Indexmerkmale werden aus den Textkonfigurationseinstellungen übernommen.

```
db2tx "ENABLE TEXT COLUMN    db2tx.sample    comment
                   HANDLE    commenthandle
                   CCSID     819
                   LANGUAGE  uk_english
                   FORMAT    rft"
```

Im nächsten Beispiel werden die Werte für die Merkmale des Indexes, der für die Spalte COMMENT erstellt wird, explizit gesetzt. Das Beispiel setzt den Index und das Indexverzeichnis und setzt die Indexaktualisierungsfrequenz so, daß der Index um 12:00 oder 15:00 Uhr von Montag bis Freitag aktualisiert wird, wenn mindestens 100 Textdokumente in der Warteschlange stehen. Standardwerte für die Textdaten werden aus den Textkonfigurationseinstellungen übernommen.

Für einen UNIX-Server:

```
db2tx "ENABLE TEXT COLUMN      db2tx.sample    comment
                   HANDLE      commenthandle
                   INDEXTYPE   linguistic
                   UPDATEFREQ  min(100) d(1,2,3,4,5) h(12,15) m(00)
                   UPDATEINDEX UPDATE
                   DIRECTORY   DB2TX_INSTOWNER/db2tx/indexes"
```

Für einen OS/2- oder Windows NT-Server:

```
db2tx "ENABLE TEXT COLUMN      db2tx.sample    comment
                   HANDLE      commenthandle
                   INDEXTYPE   linguistic
                   UPDATEFREQ  min(100) d(1,2,3,4,5) h(12,15) m(00)
                   UPDATEINDEX UPDATE
                   DIRECTORY   \dmb\db2tx\indexes"
```

Textdokumente für die Suche vorbereiten

Textspalten für einen nicht unterstützten Datentyp aktivieren

Textspalten müssen vom Typ CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB sein, um vom Text Extender aktiviert werden zu können. Wenn sich die Dokumente in einer Spalte eines anderen Typs befinden, z.&rbl;B. ein benutzerdefinierter eindeutiger Typ (UDT), müssen Sie eine benutzerdefinierte Funktion liefern, die den Benutzertyp als Eingabe nimmt und CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB als Ausgabetyp liefert.

Verwenden Sie das Schlüsselwort FUNCTION in ENABLE TEXT COLUMN, um den Namen dieser Funktion anzugeben.

Beispiel: Sie wollen komprimierten Text in einer Tabelle speichern.

1. Erstellen Sie einen UDT für den Text:

```
db2 CREATE DISTINCT TYPE COMPRESSED_TEXT AS CLOB(1M)
```

2. Erstellen Sie eine Tabelle und fügen Sie den Text darin ein:

```
db2 CREATE TABLE MYTEXT (author VARCHAR(50),
                         text   COMPRESSED_TEXT)
db2 INSERT ...
```

Führen Sie folgende Schritte aus, um die Textspalte für die Verwendung durch den Text Extender zu aktivieren:

1. Erstellen Sie eine UDF mit beispielsweise dem Namen UNCOMPRESS, die einen Wert vom Typ COMPRESSED_TEXT empfängt und den entsprechenden dekomprimierten Text beispielsweise als einen CLOB(10M)-Wert zurückgibt.
2. Aktivieren Sie die Textspalte mit Hilfe des Schlüsselworts FUNCTION, um die UDF UNCOMPRESS zu identifizieren:

```
db2tx "ENABLE TEXT COLUMN MYTABLE  text
                          FUNCTION uncompress
                          HANDLE   handle
                          ..."
```

Indizes in einer Umgebung mit mehreren Knoten verwalten

Wenn die Textspalte, die Sie aktivieren, zu einer Tabelle gehört, die Teil einer Knotengruppe mit mehreren Knoten ist, muß das Indexverzeichnis, das Sie angeben, auf allen physischen Knoten verfügbar sein. Wenn Sie das Standardverzeichnis verwenden, das in der Textkonfiguration angegeben ist, stellen Sie sicher, daß der Pfad auf allen Knoten der Knotengruppe verfügbar ist. Wenn dies nicht passend ist, können Sie für jeden Knoten einen spezifischen Pfad im Befehl ENABLE TEXT COLUMN angeben.

Wenn Sie die Knotenkonfiguration einer Knotengruppe ändern, die eine für den Text Extender aktivierte Tabelle enthält, müssen Sie die Tabelle erneut indexieren. Sie können überprüfen, ob die Knotenkonfiguration geändert wurde, indem Sie den Befehl `TXNCHECK` verwenden.

Hinweis

Wenn Sie DB2 UDB Extended Enterprise Edition verwenden und Tabellen haben, die über mehrere Knoten partitioniert sind, kann es zu Unterschieden bei den Systemzeiten der Knoten kommen (wenn einer der Knoten beispielsweise auf Sommerzeit umgestellt ist). Um Fehler zu verhindern, stellen Sie sicher, daß sich die Unterschiede zwischen den Systemzeiteinstellungen der Knoten innerhalb des Wertes für die DB2-Einstellung MAX_TIME_DIFF befinden. Verwenden Sie bei Windows NT den Befehl NET TIME , um die Systemzeiteinstellungen zu synchronisieren.

Die Vorteile der Struktur eines Dokument nutzen

Ein strukturiertes Dokument ist ein Dokument, das Abschnitte, z.&rbl;B. den Dokumenttitel, Autor und das Thema, enthält. Sie können den Bereich einer Suche auf einen bestimmten Abschnitt in Dokumenten begrenzen.

Drei Arten von strukturierten Dokumenten werden unterstützt: unstrukturierte Dateien, HTML-Dokumente und XML-Dokumente. Eine unstrukturierte Datei mit formatierten Abschnitten könnte wie folgt aussehen:

```
<title>IBM Dictionary of Computing
<author>McDaniel, George
<subject>Computers, Reference, ...
```

Ein XML-Dokument könnte wie folgt aussehen:

```
<?xml version="1.0"?>
<!DOCTYPE LETTER SYSTEM "letter.dtd">

<LETTER>
 <HEADER> This tag has been skipped in the definition, so this text will
          be added to the section named LETTER.
 </HEADER>
 <ADDRESS>
    Text will added to section named LETTER/address.
  <CITY>
    Text will added to section named LETTER/address/City.
  </CITY>
 </ADDRESS>
 <CONTENT>
   Text will be added section named LETTER/Content.
  <NOSECTION>  Text will be added to section named LETTER/Content since element
               "NOSECTION" was not defined.
  </NOSECTION>
  <GREETINGS>
   Text will be added section named LETTER/Content/Greetings.
  </GREETINGS>
 </CONTENT>
</LETTER>
```

Ein dualer Index kann **nicht** verwendet werden, um die Vorteile der Dokumentstruktur zu nutzen.

Damit der Text Extender solche Abschnitte beim Indexieren beachtet, müssen Sie ein *Dokumentmodell* erstellen, das beschreibende Abschnittsnamen Ihrer Wahl für die Verwendung in Abfragen für diesen Abschnitt sowie die Formatierungsbefehle enthält, die diese Abschnitte identifizieren. Häufig werden Sie einen Abschnittsnamen angeben, der dem Befehlsnamen entspricht.

```
Abschnitt Befehl

Title      title
Author     author
Subject    subject
```

Die Dokumentmodelldatei. Sie beschreiben das Dokumentmodell in einer Dokumentmodelldatei mit der Erweiterung *ini*. Befehle, die nicht in der Dokumentmodelldatei definiert sind, werden gemäß dem Indextyp indexiert. Es gibt eine Dokumentmodelldatei für jedes Text Extender-Server-Exemplar. Wie Sie im folgenden Beispiel ersehen können, kann eine Dokumentmodelldatei mehr als ein Dokumentmodell enthalten.

```
;Comments must be preceded by a semicolon

[MODELS]
modelname=sample
modelname=play

[sample]
title=title
author=author
subject=subject
content=content

[play]
play=play
author=author
title=title
scene=scene
```

Textdokumente für die Suche vorbereiten

! Für XML-Dokumente muß der in der Dokumentmodelldatei definierte Modellname mit
! der Abschnittsnamenskennung und dem Abschnittsnamensbefehl des Stammelements
! übereinstimmen. Bei den Namen wird die Groß-/Kleinschreibung beachtet.

! Für abschnittsfähige Indizes werden gut gestaltete XML-Dokumente, die Stammele-
! mente enthalten, die mit dem Namen eines gültigen Dokumentmodells übereinstimmen,
! gemäß der Definition dieses Dokuments indexiert. Wenn kein übereinstimmendes
! Dokumentmodell gefunden wird, wird eine Fehlernachricht ausgegeben und das Doku-
! ment wird nicht indexiert. Für Indizes, die nicht abschnittsfähig sind, werden zumindest
! alle gut gestalteten XML-Dokumente indexiert und die Abschnittsnamensbefehle werden
! ignoriert.

! Hier ist ein Beispiel einer Modelldefinition für Dokumente im XML-Format. Es enthält
! *verschachtelte* Abschnitte, d.&rbl;h. Abschnitte innerhalb anderer Abschnitte.

```
! ; list of document models
! [MODELS]
! modelname = LETTER
! ;  sample for XML document model definition
! ;  left side = section name identifier encoding whole path
! ;  right side = section name tags specifying tag for each element
! ;               of the path through the tree down to specified
! ;                  node. Tag delimiter is '/'.
! [LETTER]
! LETTER = LETTER
! LETTER/address = LETTER/ADDRESS
! LETTER/Content = LETTER/CONTENT
! LETTER/address/City = LETTER/ADDRESS/CITY
! LETTER/Content/Greetings = LETTER/CONTENT/GREETINGS
```

! Der Modelleintrag LETTER zeigt, wie Abschnitte gestaltet werden, die im Index für
! Dokumente im XML-Format gespeichert werden sollen. Informationen zur Modellbaum-
! struktur werden auf der rechten Seite der Abschnittseinträge geliefert: jede Zeichenfolge
! gibt ein XML-Element an, der Begrenzer "/" gibt die Elementhierarchie an. Die in der
! Konfigurationsdatei angegebene Definition muß eine Untermenge des Modells sein, wie
! in der Dokumentartdefinitionsdatei (DTD-Datei) für ein bestimmtes Dokument definiert.

Textdokumente für die Suche vorbereiten

Der erste Abschnittseintrag für ein Modell, das für XML-Dokumente verwendet werden soll, muß für das Stammelement gelten. Die vollständige Hierarchie des Modells muß definiert werden, bis zum letzten Element, das als Abschnitt definiert ist. Die Reihenfolge muß nach dem Prinzip "höher zuerst" festgelegt sein. Das heißt, alle Elemente der Ebene n der Modellhierarchie, die für die Indexierung verwendet werden, müssen vor Elementen der Ebene n+1 deklariert werden. Sie können den Eintrag "LETTER/Content = LETTER/CONTENT" nicht auslassen, wenn Sie "LETTER/Content/Greetings = LETTER/CONTENT/GREETINGS" als Abschnitt im Index verwenden wollen. Sie können jedoch ganze Zweige der Hierarchie auslassen; Text, der in diesen Zweigen enthalten ist, wird so behandelt, als gehöre er zum Stammelement.

Einschränkungen: Die Suche in verschachtelten Abschnitten ist nur in XML-Dokumenten möglich. Bei Ngram-Indizes kann nur der Name des Abschnitts gesucht werden und dieser darf kein verschachtelter Abschnitt sein.

Wenn ein Server-Exemplar erstellt wird, wird ein Beispiel einer Dokumentmodelldatei, *desmodel.ini*, im Unterverzeichnis des Server-Exemplars erstellt. Im selben Verzeichnis befindet sich auch ein Beispiel einer Dokumentmodelldatei für HTML-Dokumente. HTML-Dokumente können keine verschachtelten Abschnitte enthalten.

Zum Aktivieren der Abschnittsunterstützung definieren Sie Dokumentmodelle in der Datei *desmodel.ini* und aktivieren Sie unter Verwendung von `INDEXPROPERTY SECTIONS_ENABLED` die Textspalte, die die Dokumente enthält. Einzelheiten finden Sie in „ENABLE TEXT COLUMN“ auf Seite 139.

Hinweise

- Bei Modellnamen und Abschnittsnamen muß die Groß-/Kleinschreibung beachtet werden und nur die Zeichen A - Z, a - z, 0 - 9 und / sind zulässig.
- Kommentare können hinzugefügt werden; ihnen muß ein Semikolon vorangestellt werden.

Textdokumente für die Suche vorbereiten

Externe Textdateien aktivieren

Zusammenfassung

Zeitpunkt Wahlfrei. Einmal für jede Tabelle, die den zu durchsuchenden externen Dateien zugeordnet ist.

Befehl ENABLE TEXT FILES

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Der Text Extender kann nicht nur in Text suchen, der in DB2 UDB-Tabellen gespeichert ist, sondern auch in Textdokumenten, die in Dateien gespeichert sind. Dieser Vorbereitungsschritt ist erforderlich, wenn Sie nach Text in externen Dateien suchen wollen. Die Tabelle, die den externen Textdateien zugeordnet ist, muß nicht durch den Befehl ENABLE TEXT TABLE aktiviert worden sein.

Ein Index und eine Protokolltabelle werden erstellt und die Dokumentinformationen werden gesetzt, genau wie in „Eine Textspalte aktivieren“ auf Seite 60 beschrieben.

Hinweise

1. Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 63.
2. Wenn Sie eine partitionierte Tabelle aktivieren, die eine Knotengruppe verwendet, die mehrere physische Knoten hat, stellen Sie sicher, daß die externen Dateien, auf die Sie in den Tabellenspalten verweisen, sich auf dem Knoten befinden, auf dem sich die entsprechende Tabellenpartition befindet.

Eine Kennungsspalte vom Typ DB2TEXTFH wird zu einer bestehenden DB2 UDB-Tabelle hinzugefügt. Die Kennungsspalte hält die Referenzen für die externen Dateien, wobei jede Kennung die Index- und Dokumentinformationen (CCSID, Format und Sprache) enthält.

In „Kennungen für externe Dateien“ auf Seite 99 finden Sie dazu eine Beschreibung.

Sie können zusätzliche Parameter, wie z.&rbl;B. die Standardindexmerkmale, genau wie beim Aktivieren einer Textspalte angeben.

Nachdem der Index erstellt wurde, können Sie die externen Dateien verschieben oder löschen. Sie können die Datei dennoch durchsuchen. Sie können neue Zeilen in die Tabelle einfügen und den Befehl UPDATE INDEX verwenden, um den Index mit den neuen Dateireferenzen zu aktualisieren.

Textdokumente für die Suche vorbereiten

Wenn Sie eine Tabelle aktivieren, die eine Knotengruppe mit mehreren physischen Knoten verwendet, stellen Sie sicher, daß die externen Dateien, auf die Sie in den Spalten Ihrer Tabelle verweisen, sich auf dem Knoten befinden, auf dem sich die Tabellenpartition befindet.

Beispiele

1. Erstellen Sie eine Tabelle DB2TX.EXTFILE, die mindestens eine Spalte hat oder eine bestehende Tabelle verwendet.
2. Fügen Sie die Kennungsspalte FILEHANDLES zur Tabelle DB2TX.EXTFILE hinzu:

```
db2tx "ENABLE TEXT FILES  db2tx.extfile
              HANDLE      filehandles
              INDEXTYPE   linguistic
              UPDATEFREQ  min(100) d(1,2,3,4,5) h(12,15) m(00)
              UPDATEINDEX NOUPDATE
              DIRECTORY   \any\db2tx\indices"
```

3. Initialisieren Sie die Kennung
   - für jede Zeile in einer neuen Tabelle:

```
db2 INSERT INTO db2tx.EXTFILE (FILEHANDLES)
VALUES (db2tx.INIT_TEXT_HANDLE
           (850, 'TDS' 'US_ENGLISH',
           'd:\dmb\samples\tx')
```

   - für eine bestehende Tabelle und aktualisieren Sie die Kennungsspalten, um die Referenz auf eine externe Datei wiederzugeben, indem Sie den Namen der externen Datei angeben:

```
db2 UPDATE db2tx.EXTFILE
SET FILEHANDLES = db2tx.file(FILEHANDLES,'d:\dmb\samples\tx')
WHERE DOCID = 'doc1'
```

Hinweis

Verwenden Sie nicht INIT_TEXT_HANDLE zum Aktualisieren von Kennungsspalten, die auf externe Dateien verweisen.

4. Aktualisieren Sie den Index:

```
db2tx "UPDATE INDEX  db2tx.extfile
            HANDLE  filehandles"
```

Den Textvorbereitungsprozeß umkehren

Die Verwaltungssitzung beenden

Sie haben jetzt die Schritte abgeschlossen, um Ihre Textdokumente für die Suche vorzubereiten.

Wenn Sie beim Aktivieren der Textspalte die Option NOUPDATE für das Schlüsselwort UPDATEINDEX angegeben haben, indexiert der Text Extender den Text nicht sofort, sondern wartet auf die nächste regelmäßige Indexierung. Um den Index sofort zu aktualisieren, gehen Sie zu „Einen Index aktualisieren“ auf Seite 77.

Wenn das Indexieren der Dokumente beendet ist, können Sie beginnen, Informationen abzurufen, wie in Kapitel 5, „Mit Hilfe von Text Extender-UDFs suchen“ auf Seite 95, beschrieben.

Geben Sie QUIT ein, um den Text Extender-Befehlsprozessor zu beenden.

Hinweis

Verwenden Sie GET INDEX STATUS, um festzustellen, wann das Indexieren beendet ist.

Den Textvorbereitungsprozeß umkehren

Wenn Text für die Verwendung durch den Text Extender vorbereitet wird, werden bestimmte administrative Änderungen vorgenommen. In diesem Abschnitt werden Funktionen beschrieben, die Sie beim Umkehren dieses Prozesses unterstützen.

Eine Textspalte inaktivieren

Zusammenfassung

Zeitpunkt Wenn Sie nicht mehr beabsichtigen, Textsuchen in einer Textspalte durchzuführen.

Befehl DISABLE TEXT COLUMN

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Beispiel:

```
db2tx "DISABLE TEXT COLUMN db2tx.sample
                HANDLE   commenthandle"
```

Wenn Sie eine Textspalte inaktivieren, passiert folgendes:

- Wenn es sich um eine Mehrindextabelle handelt, d.&rbl;h., die Spalte hat einen eigenen Textindex und eine eigene Protokolltabelle, werden der Index, die Protokolltabelle und die Protokolltabellenauslöser gelöscht.
- Wenn es sich um eine Tabelle mit einem gemeinsam benutzten Index handelt, d.&rbl;h., es gibt einen Index, der von allen Textspalte gemeinsam benutzt wird, werden die Begriffe für die Dokumente dieser Spalte aus dem gemeinsamen Index gelöscht. Wenn dies die einzig verbliebene aktivierte Textspalte in der Tabelle ist, werden der Index, die Protokolltabelle und die Protokolltabellenauslöser gelöscht.

Textdateien inaktivieren

Zusammenfassung

Zeitpunkt Wenn Sie nicht mehr beabsichtigen, Textsuchen in einer Gruppen von externen Textdateien durchzuführen.

Befehl DISABLE TEXT FILES

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Beispiel:

```
db2tx "DISABLE TEXT FILES db2tx.sample
               HANDLE   commenthandle"
```

Wenn Sie externe Textdateien inaktivieren, passiert folgendes:

- Der Index für die entsprechende Kennungsspalte wird gelöscht.
- Die Protokolltabelle und die Auslöser werden gelöscht.

Den Textvorbereitungsprozeß umkehren

Eine Texttabelle inaktivieren

Zusammenfassung

Zeitpunkt Wenn Sie nicht mehr beabsichtigen, Textsuchen in einer Texttabelle durchzuführen.

Befehl DISABLE TEXT TABLE

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Beispiel:

```
db2tx "DISABLE TEXT TABLE db2tx.sample"
```

Wenn Sie eine Texttabelle inaktivieren, passiert folgendes:

- Wenn ein gemeinsamer Index für die Textspalten in der Tabelle existiert, wird dieser Index gelöscht. Wenn statt dessen individuelle Indizes für jede Textspalte existieren, werden *alle* Indizes für die Textspalten gelöscht.
- Die gemeinsame Protokolltabelle, die verwendet wird, um automatisch aufzuzeichnen, welche Textdokumente indexiert werden müssen, wird gelöscht. Wenn statt dessen individuelle Protokolltabellen für jede Textspalte existieren, werden alle Protokolltabellen gelöscht.
- Die Auslöser, die zur Verwaltung der Protokolltabellen verwendet werden, werden gelöscht.
- Der Inhalt der Kennungsspalten wird auf Null gesetzt.

Eine Datenbank inaktivieren

Zusammenfassung

Zeitpunkt Wenn Sie nicht mehr beabsichtigen, Textsuchen in dieser Datenbank durchzuführen.

Befehl DISABLE DATABASE

Berechtigung SYSADM oder DBADM für die Datenbank

Um die verbundene Datenbank zu inaktivieren, geben Sie folgendes ein:

```
db2tx "DISABLE DATABASE"
```

Wenn Sie eine Datenbank inaktivieren, werden folgende Objekte gelöscht:

- Die Text Extender-Katalogsicht, die beim Aktivieren der Datenbank erstellt wurde.
- Die Deklaration der benutzerdefinierten Funktionen (UDFs) des Text Extender und die benutzerdefinierten eindeutigen Typen (UDTs) für diese Datenbank.
- Alle Indizes, die zu einer der Texttabellen oder Textspalten dieser Datenbank gehören.
- Die Protokolltabellen, die verwendet werden, um automatisch aufzuzeichnen, welche Textdokumente indexiert werden müssen, und die Auslöser, die zu deren Verwaltung verwendet werden.

Da Kennungsspalten nicht gelöscht werden können und die Kennungsspalte vom eindeutigen Typ ist, können einige eindeutige Typen nicht gelöscht werden.

Textindizes verwalten

Dies sind die Verwaltungs-Tasks:

- Einen Index aktualisieren
- Die Einstellungen eines Indexes ändern
- Den Status eines Indexes zurücksetzen
- Indexereignisse löschen
- ! Einen Index reorganisieren

Sie können diese Tasks zu einem beliebigen Zeitpunkt und in einer beliebigen Reihenfolge ausführen.

Einen Index aktualisieren

Zusammenfassung

Zeitpunkt Wenn ein Index sofort aktualisiert werden muß, ohne auf die regelmäßige Indexierung zu warten. (Informationen zur regelmäßigen Indexierung finden Sie in „Eine Textspalte aktivieren“ auf Seite 60 und „Die Einstellungen eines Indexes ändern“ auf Seite 79.)

Befehl UPDATE INDEX

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Dieses Beispiel aktualisiert den Index für eine Tabelle mit einem gemeinsam benutzten Index:

```
db2tx "UPDATE INDEX db2tx.sample"
```

Dieses Beispiel aktualisiert den Index für eine Spalte einer Mehrindextabelle:

```
db2tx "UPDATE INDEX db2tx.sample HANDLE commenthandle"
```

Verwenden Sie diesen Befehl, um den Index sofort zu aktualisieren, ohne darauf zu warten, daß die nächste regelmäßige Indexierung automatisch stattfindet. Dies ist sinnvoll, wenn Sie mehrere Textdokumente zu einer Datenbank hinzugefügt haben und sie sofort durchsuchen wollen.

Der Text Extender indexiert die Textdokumente in dieser Spalte (bzw. alle Spalten in der Tabelle), die eingefügt oder geändert wurden, und löscht die Begriffe aus dem Index, die aus Dokumenten stammen, die gelöscht wurden. Die zum Index zugehörige Protokolltabelle enthält Informationen dazu, welche Dokumente eingefügt, aktualisiert und gelöscht wurden.

Einen Index für externe Dateien aktualisieren

Eine Protokolltabelle enthält nicht automatisch Informationen zu Änderungen an externen Dateien, die Sie möglicherweise indexiert haben (siehe „Externe Textdateien aktivieren“ auf Seite 72). Zu solchen Änderungen kann das Ersetzen eines Dokuments durch eine neuere Version mit demselben absoluten Pfadnamen gehören. Aktualisierungen an externen Dateien können nicht durch den Text Extender in Protokolltabellen überwacht werden, da diese Aktualisierungen nicht innerhalb des Bereichs von DB2 UDB auftreten.

Damit die Aktualisierungen an externen Dateien in einem Text Extender-Index wiedergegeben werden, können Sie folgende Schritte ausführen:

1. Erzwingen Sie, daß ein “Änderungseintrag” in die Protokolltabelle gestellt wird, indem Sie eine UPDATE-Anweisung für die entsprechende Kennungsspalte ausgeben, die effektiv nichts tut:

   ```
   UPDATE tabelle
   SET    dateikennungsspalte = dateikennungsspalte
   WHERE  DB2TX.FILE(dateikennungsspalte) = dateiname
   ```

   Dabei ist *dateiname* der absolute Pfadname der externen Datei, die aktualisiert wurde.

2. Führen Sie UPDATE INDEX aus, um den Index auf den neuesten Stand zu bringen, einschließlich der Änderungen an der externen Datei.

Die Einstellungen eines Indexes ändern

Zusammenfassung

Zeitpunkt Wenn die Aktualisierungsfrequenz eines Indexes geändert werden muß.

Befehl CHANGE INDEX SETTINGS

Berechtigung ALTER, SELECT, UPDATE für die Tabelle

Verwenden Sie diesen Befehl, um die Aktualisierungsfrequenz eines Indexes zu ändern.

Aktualisierungsfrequenz

Weitere Informationen finden Sie in „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262. Wenn Sie keine Aktualisierungsfrequenz angeben, bleiben die aktuellen Einstellungen unverändert. Verwenden Sie das Schlüsselwort UPDATEINDEX, um festzulegen, ob das Indexieren der Textdokumente sofort beginnen soll oder bei der nächsten geplanten regelmäßigen Indexierung. Wenn Sie dieses Schlüsselwort nicht verwenden, bleibt die aktuelle Einstellung unverändert.

Beispiele

Um die Aktualisierungsfrequenz für den Index zu ändern, so daß er um 12:00 oder 15:00 Uhr von Montag bis Freitag aktualisiert wird, wenn mindestens 100 Textdokumente in der Warteschlange stehen, geben Sie folgendes ein:

```
db2tx "CHANGE INDEX SETTINGS db2tx.sample
                  HANDLE     commenthandle
                  UPDATEFREQ min(100) d(1,2,3,4,5) h(12,15) m(00)"
```

Um das regelmäßige Aktualisieren eines Indexes zu stoppen, geben Sie folgendes ein:

```
db2tx "CHANGE INDEX SETTINGS db2tx.sample
                  HANDLE     commenthandle
                  UPDATEFREQ none"
```

Den Indexstatus zurücksetzen

Zusammenfassung

Zeitpunkt Wenn ein Index nicht länger durchsucht oder aktualisiert werden kann.

Befehl RESET INDEX STATUS

Berechtigung Keine

Textindizes verwalten

Es können Situationen auftreten, die verhindern, daß Sie in einem Index suchen oder einen Index aktualisieren. In „Den Status eines Indexes anzeigen“ auf Seite 84 wird beschrieben, wie Sie feststellen können, ob eines dieser Ereignisse aufgetreten ist. RESET INDEX STATUS reaktiviert den Index, so daß Sie ihn erneut verwenden können.

Dieses Beispiel setzt den Indexstatus für den Index einer Tabelle mit einem gemeinsam benutzten Index zurück:

```
db2tx "RESET INDEX STATUS db2tx.sample"
```

Mit dieser Syntax können Sie auch den Indexstatus für eine bestimmte Textspalte zurücksetzen. Dieses Beispiel setzt den Indexstatus für den Index einer Mehrindextabellenspalte zurück:

```
db2tx "RESET INDEX STATUS db2tx.sample HANDLE commenthandle"
```

Indexereignisse löschen

Zusammenfassung

Zeitpunkt Wenn Sie die Nachrichten in der Protokolltabelle eines Indexes nicht mehr benötigen.

Befehl DELETE INDEX EVENTS

Berechtigung Keine

Wenn etwas verhindert, daß Sie in einem Index suchen oder einen Index aktualisieren oder wenn ein Dokument nicht indexiert werden kann, ist dies als *Indexierungsereignis* bekannt. Informationen zu Indexierungsereignissen werden in der Protokolltabelle des Indexes gespeichert. Sie können Sie bei der Bestimmung der Ursache für den Fehler unterstützen. Wenn Sie diese Nachrichten nicht mehr benötigen, können Sie sie löschen.

Dieses Beispiel löscht Nachrichten aus dem Index einer Tabelle mit einem gemeinsam benutzten Index:

```
db2tx "DELETE INDEX EVENTS db2tx.sample"
```

Mit dieser Syntax können Sie auch die Indexierungsereignisse für eine bestimmte Textspalte löschen. Dieses Beispiel löscht die Nachrichten für den Index einer Mehrindextabellenspalte:

```
db2tx "DELETE INDEX EVENTS db2tx.sample HANDLE commenthandle"
```

Einen Index reorganisieren

Zusammenfassung

Zeitpunkt Wenn GET INDEX STATUS angibt, daß ein Index manuell reorganisiert werden sollte.

Befehl REORGANIZE INDEX

Berechtigung Keine

Wenn eine Textspalte oft aktualisiert wird, wird die Suche im Index ineffizient. Um die Suche wieder effizient zu machen, muß der Index *reorganisiert* werden. Obwohl der Text Extender erkennt, wenn ein Index reorganisiert werden muß und dies automatisch im Hintergrund durchführt, können möglicherweise Situationen auftreten, in denen es erforderlich ist, einen Index manuell unter Verwendung von REORGANIZE INDEX zu reorganisieren. Sie können den Befehl GET INDEX STATUS verwenden, um zu ermitteln, ob ein Index reorganisiert werden muß.

Obwohl während der Ausführung von REORGANIZE INDEX Suchen durchgeführt werden können, ist dies für Indexaktualisierungen nicht möglich.

Dieses Beispiel reorganisiert den Index einer Tabelle mit einem gemeinsam benutzten Index:

```
db2tx "REORGANIZE INDEX db2tx.sample"
```

Dieses Beispiel reorganisiert den Index einer Mehrindextabellenspalte:

```
db2tx "REORGANIZE INDEX db2tx.sample HANDLE commenthandle"
```

Nützliche Informationen abrufen

In diesem Abschnitt werden die Verwaltungsbefehle beschrieben, die für die Anzeige von Informationen zu folgenden Themen verwendet werden:

- Der aktivierte Status von Datenbanken, Tabellen, Spalten und Dateien
- Die Einstellungen der Umgebungsvariablen
- Die Textkonfigurationseinstellungen
- Der Indexstatus
- Die Fehlerereignisse
- Die Indexeinstellungen
- Die Texteinstellungen für eine Spalte

Informationen zum aktivierten Status anzeigen

Zusammenfassung

Zeitpunkt	Wenn Sie Informationen zu dem aktivierten Status von Datenbanken, Tabellen, Textspalten oder externen Dateien benötigen.
Befehl	GET STATUS
Berechtigung	Keine

Geben Sie folgenden Befehl ein:

```
db2tx "GET STATUS"
```

Hier ist ein Beispiel der Ausgabe, die von GET STATUS angezeigt wird. Sie zeigt den aktivierten Status der Datenbank und aller aktivierten Tabellen, Textspalten oder Textdateien, die sie enthält.

```
Die Datenbank ist für DB2 Text Extender aktiviert

  Tabelle DB2TX.MYTABLE ist als Tabelle mit gemeinsam benutzten Index aktiviert

  Tabelle DB2TX.SAMPLE ist als Tabelle mit gemeinsam benutzten Index aktiviert
     TextColumnName     HandleColumnName
     --------------     ----------------
     COMMENT            COMMENTHANDLE


  Tabelle DB2TX.TEST ist als Tabelle mit Mehrindex aktiviert
     TextColumnName     HandleColumnName
     --------------     ----------------
     ABSTRACT1          ABSTRACT1HANDLE
     ABSTRACT2          ABSTRACT2HANDLE
```

Die Einstellungen der Umgebungsvariablen anzeigen

Zusammenfassung

Zeitpunkt	Wenn Sie Informationen zu den Einstellungen der Umgebungsvariablen benötigen.
Befehl	GET ENVIRONMENT
Berechtigung	Keine

Geben Sie folgenden Befehl ein:

```
db2tx "GET ENVIRONMENT"
```

Hier ist ein Beispiel der Ausgabe, die von GET ENVIRONMENT angezeigt wird. Sie zeigt die aktuellen Einstellungen der Text Extender-Umgebungsvariablen.

```
Aktuelles Datenbankmanagerexemplar (DB2INSTANCE) = db2
Standarddatenbankname                 (DB2DBDFT) = my_database
Exemplarname                    (DB2TX_INSTOWNER) = user1
Exemplarverzeichnis     (DB2TX_INSTOWNERHOMEDIR) = /usr/instance1
```

Nützliche Informationen abrufen

Die Textkonfigurationseinstellungen anzeigen

Zusammenfassung

Zeitpunkt Wenn Sie die Standardeinstellungen für Text, Index und Verarbeitungsinformationen benötigen.

Befehl GET TEXT CONFIGURATION

Berechtigung Keine

Diese Einstellungen sind in „Textkonfigurationseinstellungen“ auf Seite 254 beschrieben. Um sie zu ändern, siehe „Die Textkonfiguration ändern“ auf Seite 50.

Um die Textkonfiguration anzuzeigen, geben Sie folgendes ein:

```
db2tx "GET TEXT CFG"
```

Hier ist ein Beispiel der Ausgabe, die von GET TEXT CONFIGURATION angezeigt wird. Sie zeigt die aktuellen Textkonfigurationseinstellungen.

```
ID für codierten Zeichensatz (CCSID) = 850
Sprache                   (LANGUAGE) = US_ENGLISH
Format                      (FORMAT) = TDS

Indextyp                 (INDEXTYPE) = LINGUISTIC
Aktualisierungsfrequenz (UPDATEFREQ) = NONE
Indexverzeichnis         (DIRECTORY) = user1/db2tx/indexes

Indexaktualisierung   (UPDATEINDEX) = UPDATE
COMMIT ausführen nach (COMMITCOUNT) = 10 000
Tabellenbereich        (TABLESPACE) = TXLOG
```

Den Status eines Indexes anzeigen

Zusammenfassung

Zeitpunkt Wenn Sie feststellen müssen, ob ein Index durchsucht oder aktualisiert werden kann.

Befehl GET INDEX STATUS

Berechtigung Keine

Es können Situationen auftreten, die verhindern, daß Sie in einem Index suchen oder einen Index aktualisieren. In solchen Situationen werden Nachrichten in der Protokolltabelle des Indexes gespeichert, die Sie bei der Bestimmung der Ursache unterstützen. Es kann also nützlich sein, den Status eines Indexes zu prüfen und zu prüfen, ob Nachrichten verfügbar sind.

Dieses Beispiel zeigt den Indexstatus für den Index einer Tabelle mit einem gemeinsam benutzten Index:

```
db2tx "GET INDEX STATUS db2tx.sample"
```

Mit dieser Syntax können Sie auch den Indexstatus für eine bestimmte Textspalte anzeigen. Dieses Beispiel ruft den Indexstatus für den Index einer Mehrindextabellenspalte ab:

```
db2tx "GET INDEX STATUS db2tx.sample HANDLE commenthandle"
```

Hier ist ein Beispiel der Ausgabe, die von GET INDEX STATUS angezeigt wird.

```
Knoten 1
Suchstatus                = Suche verfügbar
Aktualisierungstatus      = Aktualisierung verfügbar
Reorganisationsstatus     = Gestartet 13.55
Geplante Dokumente        = 0
Indexierte Dokumente      = 187000
Primärindexdokumente      = 130000
Sekundärindexdokumente    =  57000
Fehlerereignisse          = Keine Fehlerereignisse
```

Wenn der Index über mehrere Knoten verteilt ist, werden die Knoteninformationen pro Knoten angezeigt.

Suchstatus
Gibt an, ob Sie die angegebene Kennungsspalte verwenden können, um im Index zu suchen. Wenn Suchen nicht verfügbar ist, prüfen Sie den angegebenen Ursachencode nach weiteren Informationen, warum die Situation eingetreten ist. Verwenden Sie anschließend den Befehl RESET INDEX STATUS, um erneut mit dem Index arbeiten zu können. Siehe Kapitel 16, „Ursachencodes für Fehlerereignisse“ auf Seite 267.

Aktualisierungsstatus
Gibt an, ob Sie den Index für die angegebene Tabelle oder Spalte aktualisieren können. Wenn die Indexaktualisierungsfunktion nicht verfügbar ist, prüfen Sie den angegebenen Ursachencode nach weiteren Informationen, warum die Situation eingetreten ist. Verwenden Sie anschließend den Befehl RESET INDEX STATUS, um erneut mit dem Index arbeiten zu können.

Reorganisationsstatus
Gibt an, ob Sie den Index für die angegebene Tabelle oder Spalte reorganisieren können. Wenn die Funktion zum Reorganisieren nicht verfügbar ist, prüfen Sie den angegebenen Ursachencode nach weiteren Informationen, warum die Situation eingetreten ist. Ein häufiger Grund dafür, daß das Reorganisieren nicht verfügbar ist, liegt darin, daß der Index zu diesem Zeitpunkt aktualisiert wird.

Nützliche Informationen abrufen

Geplante Dokumente
Zeigt die Anzahl von Dokumenten an, die in der Warteschlange für das Indexieren (oder zum Löschen aus dem Index) aufgelistet sind.

Indexierte Dokumente
Zeigt die Anzahl von Dokumenten an, die aus der Warteschlange der geplanten Dokumente bereits indexiert wurden.

! **Primärindexdokumente**
! Zeigt die Anzahl von Dokumenten im Primärindex an.

! **Sekundärindexdokumente**
! Zeigt die Anzahl von Dokumenten im Sekundärindex an.

Fehlerereignisse
Zeigt die Anzahl von Ereignissen an, die in der Protokolltabelle des Indexes verfügbar sind. Sie können diese Informationen anzeigen, wie in „Fehlerereignisse anzeigen“ beschrieben. Wenn Sie diese Informationen nicht mehr benötigen, können Sie sie löschen, wie in „Indexereignisse löschen“ auf Seite 80 beschrieben.

Fehlerereignisse anzeigen

Wenn Probleme während der Indexierung auftreten, z.&rbl;B., wenn ein Dokument, das für die Indexierung eingeplant war, nicht gefunden werden konnte, werden diese sogenannten *Fehlerereignisse* in die Protokolltabelle des Indexes geschrieben.

Die Ereignisrückkehrcodes sind in Kapitel 16, „Ursachencodes für Fehlerereignisse“ auf Seite 267, beschrieben.

Sie können auf diese Fehlerereignisse in einer Sicht der Protokolltabelle, `db2tx.LOGIXnnnnnn`, zugreifen, wobei *IXnnnnnn* der Name des Indexes ist, der aus der Katalogsicht abrufbar ist.

Geben Sie zum Abrufen des Namens des Indexes folgendes ein:

```
DB2 SELECT TABLENAME,
           HANDLENAME,
           INDEXNAME
    FROM   DB2TX.TEXTCOLUMNS
```

Die Fehlerereignissicht hat folgendes Layout:

```
UPDATESTATUS   SMALLINT
EVENTREASON    INTEGER
EVENTMESSAGE   VARCHAR(1024)
UPDATETIME     TIMESTAMP
HANDLE         DB2TEXTH oder DB2TEXTFH
NODENUMBER     INTEGER
```

Hier ist ein Beispiel, wie auf die Informationen aus dem Indexprotokoll zugegriffen werden kann:

```
DB2 SELECT EVENTREASON,
           EVENTMESSAGE,
           UPDATETIME,
           HANDLE
    FROM   DB2TX.LOGIXNNNNNN
```

Die Indexeinstellungen anzeigen

Zusammenfassung

Zeitpunkt	Wenn Sie Informationen zu den Einstellungen eines Indexes benötigen.
Befehl	GET INDEX SETTINGS
Berechtigung	Keine

Dieses Beispiel ruft die Indexeinstellungen für den Index einer Tabelle mit einem gemeinsam benutzten Index ab:

```
db2tx "GET INDEX SETTINGS db2tx.sample"
```

Dieses Beispiel ruft die Indexeinstellungen für den Index einer Mehrindextabellenspalte ab:

```
db2tx "GET INDEX SETTINGS db2tx.sample
                HANDLE   commenthandle"
```

Wenn die Tabelle als Mehrindextabelle aktiviert ist, zeigt dieser Befehl die Indexeinstellungen aller aktivierten Textspalten in der Tabelle an. Hier ist ein Beispiel der Ausgabe, die von GET INDEX SETTINGS für eine Tabelle mit einem gemeinsam benutzten Index angezeigt wird. Die Ausgabe für eine Mehrindextabelle zeigt ähnliche Informationen für jeden Index. Mit dieser Syntax können Sie die Indexeinstellungen für eine bestimmte Textspalte anfordern.

```
Aktuelle Indexeinstellungen:

     Indextyp                (INDEXTYPE) = LINGUISTIC
     Indexaktualisierung    (UPDATEINDEX) = UPDATE
     Aktualisierungsfrequenz (UPDATEFREQ) = NONE
     Knoten 1
     Indexverzeichnis         (DIRECTORY) = user1/db2tx/indexes
```

Wenn der Index über mehrere Knoten verteilt ist, werden die Knoteninformationen für das Indexverzeichnis angezeigt.

Nützliche Informationen abrufen

Die Texteinstellungen für eine Spalte anzeigen

Zusammenfassung

Zeitpunkt Wenn Sie Informationen zu den Texteinstellungen für eine Spalte benötigen.

Befehl GET TEXT INFO

Berechtigung Keine

Dieses Beispiel ruft die Textdaten für den Index einer Tabelle mit einem gemeinsam benutzten Index ab:

```
db2tx "GET TEXT INFO db2tx.sample"
```

Dieses Beispiel ruft die Textdaten für den Index einer Mehrindextabellenspalte ab:

```
db2tx "GET TEXT INFO db2tx.sample HANDLE commenthandle"
```

Mit dieser Syntax können Sie einen Tabellennamen und den Namen einer Kennungsspalte angeben.

Wenn Sie im Befehl nur einen Tabellennamen angeben, werden die Textdaten für jede aktivierte Spalte in dieser Tabelle angezeigt. Wenn Sie außerdem einen Namen für eine Kennungsspalte angeben, werden nur die Textdaten für diese Spalte angezeigt.

Hier ist ein Beispiel dafür, was durch diesen Befehl für eine Mehrindextabelle angezeigt wird:

```
  Textinformationen für Spalte ABSTRACT1
         mit Kennungsspalte ABSTRACT1HANDLE:
  ID für codierten Zeichensatz (CCSID) = 850
  Sprache                   (LANGUAGE) = US_ENGLISH
  Format                      (FORMAT) = TDS

Textinformationen für Spalte ABSTRACT2
         mit Kennungsspalte ABSTRACT2HANDLE:
  ID für codierten Zeichensatz (CCSID) = 850
  Sprache                   (LANGUAGE) = US_ENGLISH
  Format                      (FORMAT) = TDS
```

Mit der Text Extender-Katalogsicht arbeiten

Der Text Extender erstellt und verwaltet eine Katalogsicht mit dem Namen DB2TX.TEXTINDEXES für jede Datenbank. Die Sicht wird erstellt, wenn Sie den Befehl ENABLE DATABASE ausführen. Sie enthält Informationen zu den Tabellen und Spalten, die für den Text Extender aktiviert sind.

In DB2TX.TEXTINDEXES werden immer neue Einträge erstellt, wenn eine Tabelle, eine Spalte oder externe Dateien aktiviert werden. Einträge werden immer aktualisiert, wenn die Indexeinstellungen unter Verwendung des Befehls CHANGE INDEX SETTINGS geändert werden. Einträge werden gelöscht, wenn Spalten oder Tabellen inaktiviert werden.

Daten in der Katalogsicht sind über normale SQL-Abfragefunktionen verfügbar. Sie können die Katalogsicht jedoch nicht unter Verwendung normaler SQL-Datenbearbeitungsbefehle ändern. Sie können die Katalogsicht nicht explizit erstellen oder löschen. Tabelle 5 zeigt den Inhalt der Katalogsicht.

Tabelle 5 (Teil 1 von 3). Text Extender-Katalogsicht

Spaltenname	Datentyp	Nullwerte möglich	Beschreibung
TABLESCHEMA	CHAR(8)	Nein	Das Schema der Tabelle, für die dieser Eintrag gilt.
TABLENAME	VARCHAR(18)	Nein	Der Name der Tabelle, für die dieser Eintrag gilt.
COLUMNNAME	VARCHAR(18)	Ja	Der Name einer Spalte, die innerhalb dieser Tabelle aktiviert wurde. Dieser Wert ist Null, wenn die Tabelle aktiviert wurde, aber keine Spalte aktiviert wurde.
HANDLENAME	VARCHAR(18)	Ja	Der Name einer Kennungsspalte. Dieser Wert ist Null, wenn in der Tabelle TABLESCHEMA.TABLENAME keine Spalte aktiviert ist.
INDEXNAME	CHAR(8)	Nein	Der Name des Textindexes, der während der Aktivierung der Texttabelle oder einer Textspalte erstellt wurde.
LOGTABLE	VARCHAR(18)	Nein	Der Name einer Protokolltabelle für den Index INDEXNAME. Die Tabelle DB2TX.LOGTABLE enthält Informationen, welche Textdokumente für die nächste Aktualisierung des Textindexes eingeplant sind, und Fehlerereignisse.
INDEXTYPE	VARCHAR(30)	Nein	Der Indextyp: DUAL, LINGUISTIC, PRECISE, NGRAM.

Mit der Text Extender-Katalogsicht arbeiten

Tabelle 5 (Teil 2 von 3). Text Extender-Katalogsicht

Spaltenname	Datentyp	Null-werte mög-lich	Beschreibung
MINIMUM	INTEGER	Ja	Die kleinste Anzahl von Indexaktualisierungsanforderungen, die erforderlich ist, bevor eine Indexaktualisierung ausgeführt wird. Siehe „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262. Dieser Wert ist Null, wenn die Aktualisierungsfrequenz auf NONE gesetzt ist.
DAYS	VARCHAR(15)	Ja	Die Tage, an denen eine Aktualisierung eingeplant werden soll. Siehe „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262. Dieser Wert ist Null, wenn die Aktualisierungsfrequenz auf NONE gesetzt ist.
HOURS	VARCHAR(75)	Ja	Die Stunden, zu denen eine Indexaktualisierung eingeplant werden soll. Siehe „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262. Dieser Wert ist Null, wenn die Aktualisierungsfrequenz auf NONE gesetzt ist.
MINUTES	VARCHAR(185)	Ja	Die Minuten, zu denen eine Aktualisierung eingeplant werden soll. Siehe „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262. Dieser Wert ist Null, wenn die Aktualisierungsfrequenz auf NONE gesetzt ist.
INDEXDIRECTORY	VARCHAR(254)	Nein	Der Name des Verzeichnisses, in dem der Textindex innerhalb des Dateisystems gespeichert werden soll.
UPDATEONCREATE	VARCHAR(10)	Nein	Der Wert “UPDATE” bzw. “NOUPDATE”, je nachdem, welcher Wert mit der Option UPDATEINDEX im Befehl ENABLE TEXT TABLE oder ENABLE TEXT COLUMN oder im letzten Befehl CHANGE INDEX SETTINGS angegeben wurde.
COMMONINDEX	VARCHAR(4)	Nein	“YES”, wenn die Tabelle TABLESCHEMA.TABLENAME eine Tabelle mit einem gemeinsam benutzten Index ist. “NO”, wenn die Tabelle TABLESCHEMA.TABLENAME eine Mehrindextabelle ist.
CCSID	SMALLINT	Ja	Die CCSID (ID für codierten Zeichensatz) für die Textspalte TEXTCOLUMN, die mit dem Befehl ENABLE TEXT COLUMN angegeben wurde. Dieser Wert ist Null, wenn TEXTCOLUMN Null ist.
LANGUAGE	VARCHAR(30)	Ja	Der Name des Wörterverzeichnisses, das bei der Verarbeitung der Textspalte TEXTCOLUMN verwendet wurde. Dieser Wert ist Null, wenn TEXTCOLUMN Null ist.

Tabelle 5 (Teil 3 von 3). Text Extender-Katalogsicht

Spaltenname	Datentyp	Null-werte mög-lich	Beschreibung
FORMAT	VARCHAR(30)	Ja	Das Format, das für die Textspalte TEXTCOLUMN angegeben ist. Dieser Wert ist Null, wenn TEXTCOLUMN Null ist.
FUNCTIONSCHEMA	CHAR(8)	Ja	Das Schema der Zugriffs-UDF, die mit Hilfe der Option FUNCTION im Befehl ENABLE TEXT COLUMN angegeben wurde. Dieser Wert ist Null, wenn keine Option FUNCTION angegeben ist.
FUNCTIONNAME	VARCHAR(18)	Ja	Der Name der Zugriffs-UDF, die mit Hilfe der Option FUNCTION im Befehl ENABLE TEXT COLUMN angegeben wurde. Dieser Wert ist Null, wenn keine Option FUNCTION angegeben ist.
PROTOTYPEHANDLE	VARCHAR(60)	Ja	Eine Kennung für die Verwendung in Leistungs-UDFs. Sie enthält nur den Indexnamen, der für die gesamte Textspalte gilt.
INDEXOPTION	VARCHAR(30)	Ja	Die Option, die bei der Erstellung des Indexes verwendet wurde: CASE_ENABLED.
INDEXPROPERTY	VARCHAR(30)	Ja	Das Merkmal, das bei der Erstellung des Indexes verwendet wurde: SECTIONS_ENABLED
NODENUMBER	INTEGER	Nein	Die Knotennummer der Tabellenpartition.

Die Trace-Funktion ausführen

Wenn Sie einen Fehler an einen IBM Ansprechpartner berichten müssen, werden Sie möglicherweise aufgefordert, den Trace einzuschalten, so daß Informationen in eine Datei geschrieben werden können, die zum Lokalisieren des Fehlers verwendet werden kann. Verwenden Sie die Trace-Funktion nur, wenn Sie von einem Mitarbeiter der IBM Unterstützungsfunktion oder des technischen Dienstes dazu aufgefordert werden.

Die Systemleistung wird beeinträchtigt, wenn die Trace-Funktion eingeschaltet ist, verwenden Sie sie also nur, wenn Fehlerbedingungen vorliegen.

Um die Trace-Funktion einzuschalten, geben Sie folgenden Befehl ein:

```
txtrace on optionen
```

Die Syntax und Listen der Ereignisse und Komponenten werden in „TXTRACE“ auf Seite 172 beschrieben. Eine Beschreibung weitere Optionen finden Sie ebenfalls dort.

Die Trace-Funktion ausführen

Sie können den Trace filtern, indem Sie eine “Maske” angeben, die dazu führt, daß der Trace jeden Trace-Satz auf der Basis seiner ID akzeptiert oder zurückweist. Standardmäßig wird die Trace-Funktion für alles ausgeführt.

Eine Maske besteht aus vier Teilen, die durch Punkte voneinander getrennt sind, z.&rbl;B. `2.2-6.1,3.*`, wobei die Zahlen folgende Bedeutung haben:

2 gibt den DB2 UDB Text Extender an.

2-6 schließt nur Einträge mit einer Ereignis-ID zwischen 2 und 6 ein

1,3 schließt nur Ereignisse ein, die durch die Komponenten 1 und 3 berichtet wurden

***** schließt alle Funktionen der Komponenten ein

Sie können Systemfehler unterhalb einer bestimmten Wertigkeit ausschließen und Sie können angeben, ob, wenn der Trace-Puffer voll wird, die ersten oder die letzten Datensätze beibehalten werden sollen.

Um den Fehler erneut zu erzeugen und die Trace-Informationen binär in eine Speicherauszugsdatei zu schreiben, geben Sie folgenden Befehl ein:

```
txtrace dump speicherauszugsdatei
```

Nachdem Sie die Trace-Informationen in eine Speicherauszugsdatei geschrieben haben, schalten Sie die Trace-Funktion aus:

```
txtrace off
```

Um eine formatierte Version der Speicherauszugsdatei zu erstellen, geben Sie folgendes ein:

```
txtrace format  speicherauszugsdatei formatierte_datei
```

Sie können die Trace-Informationen auch direkt aus dem gemeinsam benutzten Speicher in eine formatierte Datei schreiben, während die Trace-Funktion eingeschaltet ist:

```
txtrace format > formatierte_datei
```

Indizes und aktivierte Datenbanken sichern und zurückschreiben

Sie können aktivierte Datenbanken und Textindizes, die der Text Extender erstellt hat, sichern und zurückschreiben.

Führen Sie zum **Sichern** folgende Schritte aus:

1. Stellen Sie fest, welche Tabellen durch den Text Extender aktiviert wurden. Geben Sie dazu folgenden Befehl ein:

   ```
   db2tx "GET STATUS"
   ```

2. Stellen Sie die Namen der Indexverzeichnisse fest, die von der Datenbank verwendet werden. Geben Sie dazu folgenden Befehl ein:

   ```
   db2tx "GET INDEX SETTINGS"
   ```

3. Stoppen Sie den Text Extender-Server. Geben Sie zu den Befehl `TXSTOP` ein.
4. Sichern Sie die Indexverzeichnisse und deren Unterverzeichnisse `index` und `work`.
5. Sichern Sie die Datei `desmastr.dat`, die sich in folgendem Verzeichnis befindet:

   ```
   exemplareigner_home_verzeichnis/db2tx/txins000 (UNIX)
   oder
   DMBMMPATH\db2tx\txins000 (OS/2 und Windows NT)
   ```

   Beachten Sie, daß für partitionierte Datenbanken mehrere Exemplarverzeichnisse existieren mit den Namen `txins000, txins001, ...`. Es wäre also angebrachter, das gesamte Verzeichnis `.../db2tx` zu sichern anstelle der Dateien in den einzelnen Verzeichnissen.

6. Starten Sie den Text Extender-Server erneut:

   ```
   TXSTART
   ```

Führen Sie zum **Zurückschreiben** folgende Schritte aus:

1. Stoppen Sie den Text Extender-Server:

   ```
   TXSTOP
   ```

2. Sichern Sie die bestehende Datei `desmastr.dat`.
3. Schreiben Sie die Sicherungskopie der Datei `desmastr.dat` zurück.
4. Schreiben Sie die Sicherungskopien der Indexverzeichnisse in demselben Pfad wie zuvor zurück.
5. Starten Sie den Text Extender-Server erneut:

   ```
   TXSTART
   ```

Indizes und aktivierte Datenbanken sichern und zurückschreiben

Kapitel 5. Mit Hilfe von Text Extender-UDFs suchen

Der Text Extender liefert SQL-Funktionen, mit denen Sie Unterabfragen zur Textsuche in SQL-Abfragen einschließen können. Diese Funktionen werden zusätzlich zu den normalerweise in SQL verfügbaren Funktionen zur Verfügung gestellt. Sie sind bei DB2 als *benutzerdefinierte Funktionen* (UDFs) bekannt.

Eine Beschreibung der Syntax für UDFs finden Sie in Kapitel 9, „UDTs und UDFs“ auf Seite 179.

Bevor Sie mit der Suche beginnen, lesen Sie den Abschnitt „Indextypen“ auf Seite 13 und verwenden Sie den Befehl GET INDEX SETTINGS, um zu ermitteln, welche Indextypen dem Text zugeordnet sind, in dem Sie suchen. Eine Suche kann je nach Indextyp zu unterschiedlichen Ergebnissen führen.

! Bei den Beispielen in diesem Kapitel wird vom Indextyp LINGUISTIC ausgegangen.

In diesem Kapitel werden folgende Themen beschrieben:

- Die Beispiel-UDFs
- Die Beispieltabelle
- Kennungen für externe Dateien
- Setzen des Funktionspfads, um SQL Zugriff auf die UDFs zu geben
- Suchen nach Text unter Verwendung von CONTAINS, NO_OF_MATCHES und RANK
- Angeben von Suchargumenten in UDFs unter Verwendung der Beispiele aus CONTAINS
- Feinanpassung einer vorherigen Suche unter Verwendung von CONTAINS und REFINE
- Setzen und Extrahieren von Informationen in Kennungen unter Verwendung von INIT_TEXT_HANDLE, CCSID, FORMAT und LANGUAGE
- Verbessern der Suchleistung unter Verwendung von HANDLE_LIST, HANDLE, NO_OF_MATCHES und SEARCH_RESULT

Die Beispiel-UDFs

Der Text Extender liefert eine Eingabedatei für den Befehlszeilenprozessor mit dem Namen `txsample.udf`. Sie enthält Beispiele von Text Extender-UDFs, die mit den Beispieltabellen ausgeführt werden, die mit dem Dienstprogramm `txsample` erstellt wurden. Verwenden Sie diese Datei, um Beispiele der Syntax von Verwaltungs- und Such-UDFs und der in den Suchargumenten verwendeten Syntax anzuzeigen.

Um diese Datei auszuführen, geben Sie folgendes ein:

```
db2 -t -v -f txsample.udf
```

Die Beispieltabelle DB2TX.SAMPLE

„Eine Beispieltabelle erstellen“ auf Seite 54 beschreibt, wie die Tabelle DB2TX.SAMPLE erstellt wird. Diese Tabelle wird in vielen der hier verwendeten Beispielen benutzt. Sie können die Beispiele selbst unter Verwendung des Text Extender testen.

Ein Auszug der Tabelle DB2TX.SAMPLE wird in Tabelle 6 gezeigt.

Tabelle 6 (Teil 1 von 2). Ein Auszug aus der Beispieltabelle DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 5	RSSHERM at CHGVMIC1	LIBDB2E.A error	1995-07-25 -20.13.59	Customer is getting a 'No such file or directory' on LIBDB2E.A. It does not appear to be the same error message that relates to the asynchronous I/O driver. He is using beta 4 on 3.2.5. I have had him compare the permissions and ownership of /usr/lpp/db2_02_01/lib files with mine, and they are now the same. His .profile and ENV also look good. He has, unfortunately, COMMITTED the install. What else could be wrong.
doc 6	EDWARDSC at SYDVM1	Lowercase Userid and Password from DDCS/2	1995-07-25 -20.15.20	After rechecking, the instance where I had problems with case-sensitivity was using a DB2/2 gateway to MVS. It didn't like it when I passed a lower case userid (didn't care about passwd). Connection was only successful if I actually typed an upper case userid. So, I guess this doesn't help your situation. Sorry.
doc 7	SKY at TOROLAB4	ODBC & Stored Procedures	1995-07-25 -20.42.27	There are two sets of sample programs explaining the use of Stored Procedures using CLI (ODBC). The C file inpsrv2.c (placed on the server), and the C file inpcli2.c (placed on the client) make up the sample that demonstrates using stored procedures for input. The files outsrv2.c and outcli2.c make up the sample that demonstrates using stored procedures for output. These files are part of the .../sqllib/samples/cli files. The MAKE file will automatically build them and transfer the server file to the correct subdirectory.

Tabelle 6 (Teil 2 von 2). Ein Auszug aus der Beispieltabelle DB2TX.SAMPLE

DOCID	AUTHOR	SUBJECT	DATE	COMMENT
doc 8	ADAMACHE at TOROLAB2	DB2SYS.DLL access violation	1995-07-25 -21.13.22	Did you have a previous beta version installed? If so, did you remove it using Software Installer? Did you remove the database directories (SQLDBDIR and SQL00001, etc.) from previous beta drivers?
doc 9	ADAMACHE at TOROLAB2	CREATE DB = SYS3175: db2sysc.exe in db2eng.dll	1995-07-25 -21.40.09	Many DB2/2 beta users delete a previous beta with Software Installer, install beta 5 (or golden code now), create a database, and get: SYS3175: db2sysc.exe in db2eng.dll This happens because the directory format has changed between beta4 and beta5. Our DB2/2 installation does not migrate the sqldbdir directory between beta drivers. You should remove all occurrences of sqldbdir and sql000x directories and \sqllib\db2\sqldbdir directory. What you should do is delete the previous beta with Software Installer, remove all occurrences of sqldbdir and sql000x directories and \sqllib\db2\sqldbdir directory, and then install the new code.
doc 10	RSSHERM at CHGVMIC1	DB2/NT - SNA support	1995-07-25 -22.10.15	Will DB2/NT be able to act as both a server to CAE/WIN clients and also as a client (hopping) to DB2/6000 and/or DB2/MVS over an SNA network? The other alternative would be DRDA from DB2/NT to DB2/6000 and/or DB2/MVS - again via SNA, which I assume is supported?

Hier ist ein Teil der Tabellenstruktur, wobei die erste und die letzte Spalte gezeigt wird:

DB2TX.SAMPLE

DOCID	...	COMMENT
doc 1	...	Customer is ...
doc 2	...	After rechecking ...

Abbildung 12. Die Struktur der Tabelle DB2TX.SAMPLE

Die Spalte, die den zu durchsuchenden Text enthält, ist COMMENT. Bevor Sie den Text in dieser Spalte jedoch durchsuchen können, müssen Sie die Spalte COMMENT für die Verwendung durch den Text Extender vorbereiten, indem Sie den Befehl

Die Beispieltabelle DB2TX.SAMPLE

ENABLE TEXT COLUMN verwenden. Dies wird in „Textdokumente für die Suche vorbereiten“ auf Seite 50 beschrieben.

Nach diesem Vorbereitungsschritt sieht die Tabelle DB2TX.SAMPLE wie folgt aus:

DB2TX.SAMPLE

DOCID		COMMENT	COMMENTHANDLE
doc 1		Customer is ...	X'..handle..'
doc 2		After rechecking ...	X'..handle..'

Abbildung 13. Die Tabelle DB2TX.SAMPLE nach dem Aktivieren

Die Tabelle enthält jetzt eine zusätzlich Spalte für Kennungen, und jedes Textobjekt hat eine eindeutige Kennung, durch die es dargestellt wird.

Wenn Sie später Text in eine aktivierte Textspalte einfügen, erstellt ein Auslöser zum Einfügen eine Kennung dafür.

DB2TX.SAMPLE

DOCID		COMMENT	COMMENTHANDLE	
doc 1		Customer is ...	X'..handle..'	Kennungen, die durch ENABLE TEXT COLUMN erstellt wurden
doc 2		After rechecking ...	X'..handle..'	

Eingefügte Zeile:

doc 11		I have installed ...	X'..handle..'	Kennungen, die durch einen Einfügeauslöser erstellt wurden

Abbildung 14. Die Kennung für eine eingefügte Zeile wird von einem Auslöser erstellt

Eine Kennung enthält die folgenden Informationen:

- Eine Dokument-ID
- Den Namen und den Standort des zugeordneten Indexes
- Die Dokumentinformationen: CCSID, Format und Sprache

Die UDFs, die der Text Extender liefert, übernehmen eine Kennung als Parameter und führen das Speichern, Zugreifen, Suchen und Bearbeiten des Textes als Teil der SQL-Verarbeitung der Tabelle aus.

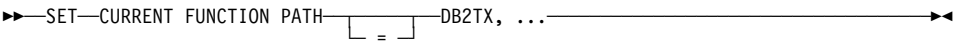
Kennungen für externe Dateien

Der Text Extender kann nicht nur in Text suchen, der in DB2 UDB-Tabellen gespeichert ist, sondern auch in Textdateien, die an anderen Orten gespeichert sind. „Externe Textdateien aktivieren“ auf Seite 72 beschreibt den Vorbereitungsschritt, der es ermöglicht, in Textdokumenten zu suchen, die nicht in DB2 UDB-Tabellen gespeichert sind. In diesem Schritt erstellt der Befehl ENABLE TEXT FILES eine Kennungsspalte vom Typ DB2TEXTFH für Kennungen für externe Dateien. Die Kennungsspalte wird zu einer bestehenden Tabelle hinzugefügt.

Sie könnten beispielsweise eine Tabelle erstellen, die Spalten für den Namen des Autors und für das Erstellungsdatum des Dokuments enthält.

Sie initialisieren die Kennungen dieser Dateien unter Verwendung von INIT_TEXT_HANDLE. Jede Kennung enthält nicht nur eine Dokument-ID, den Namen und den Standort des zugehörigen Indexes und die Dokumentinformationen (CCSID, Format und Sprache), sondern auch die Referenz auf die externe Datei.

Den aktuellen Funktionspfad setzen



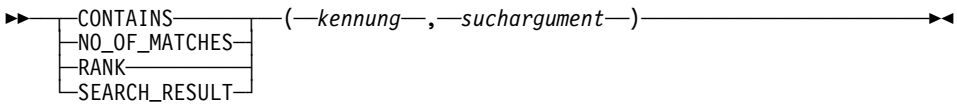
Verwenden Sie die SQL-Anweisung SET CURRENT FUNCTION PATH, um DB2TX zu den aktuellen Pfadnamen hinzuzufügen, so daß SQL die Text Extender-UDFs finden kann. Wenn Sie sich gegen diese Vorgehensweise entscheiden, können Sie die UDF-Namen explizit qualifizieren, indem Sie beispielsweise DB2TX.CONTAINS für die UDF CONTAINS eingeben.

Die Beispiele in diesem Kapitel verwenden die qualifizierte Form für Text Extender-Funktionen. Sie können die Beispielanweisungen exakt, wie sie geschrieben sind, verwenden, ohne den aktuellen Funktionspfad setzen zu müssen.

Hinweis

Denken Sie daran, den aktuellen Funktionspfad jedesmal zu setzen, wenn Sie eine Verbindung zu einer Datenbank herstellen.

Nach Text suchen



In diesem Abschnitt wird beschrieben, wie die mit dem Text Extender gelieferten UDFs verwendet werden, um in DB2-Datenbanken zu suchen, die Text enthalten. Er erläutert folgende Themen:

- Erstellen einer Abfrage
- Feststellen, wieviele Übereinstimmungen in einem Textdokument gefunden wurden
- Abrufen der Rangordnung eines gefundenen Textdokuments

Die Verwendung von SEARCH_RESULT wird in „Die Suchleistung verbessern“ auf Seite 115 beschrieben.

Jede dieser UDFs sucht im Textindex nach Vorkommen des Suchargumentes. Angenommen, es befinden sich 100 000 Textdokumente in der Tabelle, dann wird die UDF CONTAINS, RANK oder NO_OF_MATCHES 100 000 mal aufgerufen. Der Textindex wird jedoch nicht 100 000 mal durchsucht. Statt dessen wird beim ersten Aufruf der UDF eine interne Liste aller Dokumente, die den Suchbegriff enthalten, erstellt. Nachfolgende Aufrufe der UDF stellen fest, ob das betreffende Dokument sich in der Liste befindet.

Hinweis

Wenn Sie die Text Extender-UDFs für die Suche in einer Tabelle verwenden, stellen Sie sicher, daß Sie die Kennungsspalte, und nicht die Textspalte, an die UDF übergeben. Wenn Sie versuchen, in einer Textspalte zu suchen, antwortet SQL mit einer Nachricht, die angibt, daß der Datentyp falsch ist. Beispiel:

```
No function by the name "CONTAINS" having compatible
arguments was found in the function path.
```

Wenn Sie sofort nach der Ausgabe des Befehls ENABLE TEXT TABLE oder ENABLE TEXT COLUMN nach Text suchen, kann ein Fehler RC_SE_EMPTY_INDEX auftreten, der angibt, daß der vom Befehl zu erstellende Index noch nicht existiert. Die Zeit, die für die Erstellung eines Indexes erforderlich ist, hängt von verschiedenen Faktoren ab, z.&rbl;B. von der Anzahl an Dokumenten, die indexiert werden, und der Leistung des Systems, auf dem die Indexierung ausgeführt wird. Die benötigte Zeit kann von einigen Minuten bis zu einigen Stunden reichen, das Indexieren sollte also durchgeführt werden, wenn das System nur wenig belastet ist, z.&rbl;B. über Nacht.

Wenn diese Nachricht auftritt, versuchen Sie, die Suche zu einem späteren Zeitpunkt erneut durchzuführen, oder verwenden Sie den Befehl GET INDEX STATUS, um zu überprüfen, ob Indexierungsfehler aufgetreten sind.

Eine Abfrage durchführen

Dieses Beispiel zeigt, wie die Funktion CONTAINS nach Text in Dokumenten sucht, die durch eine Kennung identifiziert sind. Die Funktion gibt 1 zurück, wenn der Text das Suchargument enthält, andernfalls wird 0 zurückgegeben.

```
SELECT DATE, SUBJECT
   FROM DB2TX.SAMPLE
   WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"compress"') = 1
```

In diesem Beispiel suchen Sie nach dem Begriff `compress` in dem Text, auf den durch die Kennungen in der Spalte COMMENTHANDLE verwiesen wird. Die Kennungen in der Spalte COMMENTHANDLE geben an, wo der Text COMMENT indexiert ist.

Hinweis

Wenn Sie Kennungen in Groß-/Kleinschreibung für Tabellen oder Spalten angegeben haben, achten Sie darauf, daß diese Namen in doppelte Anführungszeichen eingeschlossen sein müssen. Beispiel:

```
SELECT DATE, SUBJECT
FROM "DB2TX.Sample"
WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"compress"') = 1
```

Wenn Sie DB2 UDB-Anweisungen `SELECT` von der Befehlszeile aus eingeben, löscht der Befehlszeilen-Parser des Betriebssystems Sonderzeichen (z.&rbl;B. doppelte Anführungszeichen) aus der Befehlszeichenfolge, so daß Sie einen umgekehrten Schrägstrich verwenden müssen, um diese Sondersymbole zu maskieren. Beispiel:

```
DB2 "SELECT DB2TX.file(COMMENTHANDLE)
    FROM DB2TX.Sample"
    WHERE DB2TX.CONTAINS (COMMENTHANDLE, '\"compress\"') = 1
```

Suchargumente angeben

Die Anzahl der gefundenen Übereinstimmungen suchen und zurückgeben

Verwenden Sie die Funktion NO_OF_MATCHES, um festzustellen, wie oft die Suchkriterien in den einzelnen Textdokumenten gefunden werden.

```
WITH TEMPTABLE(DATE, SUBJECT, MATCHES)
AS (SELECT DATE, SUBJECT,
                DB2TX.NO_OF_MATCHES(COMMENTHANDLE,'"compress"')
    FROM DB2TX.SAMPLE)
SELECT *
    FROM TEMPTABLE
         WHERE MATCHES > 0
```

NO_OF_MATCHES gibt einen Integer-Wert zurück.

Die Rangordnung eines gefundenen Textdokuments suchen und zurückgeben

RANK ist ein absoluter Wert, der angibt, wie gut das Dokument in Relation zu anderen gefundenen Dokumenten mit den Suchkriterien übereinstimmt. Der Wert gibt die Anzahl der im Dokument gefundenen Übereinstimmungen in Relation zur Größe des Dokuments an. Sie können die Rangordnung eines gefundenen Dokuments mit Hilfe der UDF RANK abrufen.

Beispiel:

```
WITH TEMPTABLE(DATE, SUBJECT, RANK)
    AS (SELECT DATE, SUBJECT,
              DB2TX.RANK(COMMENTHANDLE,'"compress"')
        FROM DB2TX.SAMPLE)
    SELECT *
        FROM TEMPTABLE
        WHERE RANK > 0
        ORDER BY RANK DESC
```

RANK gibt einen Wert DOUBLE zwischen 0 und 1 zurück.

Suchargumente angeben

Suchargumente werden in CONTAINS, NO_OF_MATCHES, RANK und HANDLE_LIST verwendet. In diesem Abschnitt wird die Funktion CONTAINS verwendet, um die unterschiedlichen Beispiele von Suchargumenten in UDFs zu zeigen.

Mehrere Begriffe suchen

Sie können mehr als einen Begriff in einem Suchargument angeben. Eine Möglichkeit, mehrere Suchbegriffe zu verbinden, ist, sie unter Verwendung von Kommas zu verbinden. Beispiel:

```
 SELECT DATE, SUBJECT
   FROM DB2TX.SAMPLE
   WHERE DB2TX.CONTAINS (COMMENTHANDLE,
     '("compress", "compiler", "pack", "zip", "compact")') = 1
```

Mit dieser Form von Suchargument wird Text gefunden, der irgend einen der Suchbegriffe enthält. In bezug auf logische Operatoren sind die Suchbegriffe durch einen OR-Operator verbunden.

Mit den Booleschen Operatoren AND und OR suchen

(Siehe auch „Mit dem Booleschen Operator NOT suchen“ auf Seite 108.)

Suchbegriffe können mit anderen Suchbegriffen kombiniert werden, indem die Booleschen Operatoren “&” (AND) und “|” (OR) verwendet werden. Beispiel:

```
SELECT DATE, SUBJECT
       FROM DB2TX.SAMPLE
       WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                             '"compress" | "compiler"') = 1
```

Sie können mehrere Begriffe unter Verwendung von Booleschen Operatoren kombinieren:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        '"compress" | "compiler" & "DB2"') = 1
```

Wenn Sie mehr als einen Booleschen Operator verwenden, wertet der Text Extender sie von links nach rechts aus, wobei allerdings der Operator für logisches Und (&) stärker bindet als der Operator für logisches Oder (|). Wenn Sie keine runden Klammern einfügen, wird beispielweise

```
"DB2" & "compiler" | "support" & "compress"
```

ausgewertet als

```
("DB2" & "compiler") | ("support" & "compress")
```

Im folgenden Beispiel müssen Sie die runden Klammern einfügen:

```
"DB2" & ("compiler" | "support") & "compress"
```

Suchargumente angeben

Wenn Sie Boolesche Operatoren mit Suchbegriffen kombinieren, die unter Verwendung des Kommatrennzeichens verkettet sind, wie in folgendem Beispiel:

```
("compress", "compiler") & "DB2"
```

wird das Komma als der Boolesche Operator OR interpretiert, wie in folgendem Beispiel:

```
("compress" | "compiler") & "DB2"
```

Nach Variationen eines Begriffs suchen

Wenn Sie einen **präzisen** Index verwenden, sucht der Text Extender nach den Begriffen exakt in der eingegebenen Schreibweise. Bei dem Begriff `media` wird nur Text gefunden, der “media” enthält. Text, der den Singular “medium” enthält, wird nicht gefunden.

Wenn Sie einen **linguistischen** Index verwenden, sucht der Text Extender auch nach Variationen der Begriffe, z.&rbl;B. Plural eines Substantivs oder eine andere Zeitform eines Verbs.

Bei dem Begriff `fahren` wird Text gefunden, der “fahren”, “fährt”, “fahrend”, “fuhr” und “gefahren” enthält.

Wenn Sie einen **dualen** Index verwenden, können Sie wählen, ob nach Wortvariationen gesucht werden soll. Beispielsweise findet die folgende Abfrage nur Vorkommen des Begriffs “utility”:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE, 'PRECISE FORM OF "utility"') = 1
```

Im Gegensatz dazu findet das folgende Beispiel Vorkommen von “utility” und “utilities”:

```
SELECT DATE, SUBJECT
       FROM DB2TX.SAMPLE
       WHERE DB2TX.CONTAINS (COMMENTHANDLE, 'STEMMED FORM OF "utility"') = 1
```

Nach Teilen eines Begriffs suchen (mit Platzhalterzeichen)

Platzhalterzeichen bieten die Möglichkeit, eine Suche flexibler zu gestalten. Sie stellen wahlfreie Zeichen am Anfang, in der Mitte oder am Ende eines Suchbegriffs dar. Sie erhöhen die Anzahl von Textdokumenten, die in einer Suche gefunden werden.

Hinweis

Wenn Sie Platzhalterzeichen verwenden, können Sie das Schlüsselwort SYNONYM FORM OF nicht verwenden. Wenn Sie einen dualen Index verwenden, wird bei der Suche mit Platzhalterzeichen die Groß-/Kleinschreibung beachtet.

Platzhalterzeichen sind besonders nützlich zum Suchen von Variationen von Begriffen, wenn ein präziser Index verwendet wird.

Wenn Sie einen linguistischen Index verwenden, werden bei Verwendung von Platzhalterzeichen viele der Variationen gefunden, die in jedem Fall gefunden worden wären.

Beachten Sie, daß Wortfragmente (Worte, die mit Platzhalterzeichen versehen sind), nicht auf eine Grundform reduziert werden werden. Wenn Sie also nach `passe%` suchen, werden Sie die Wörter “passes” oder “passed” nicht finden, da sie im Index auf ihre Grundform “pass” reduziert sind. Um sie zu finden, müssen Sie nach `pass%` suchen.

Der Text Extender verwendet zwei Platzhalterzeichen: das Unterstreichungszeichen (_) und das Prozentzeichen (%):

- % repräsentiert **eine beliebige Anzahl von willkürlichen Zeichen**. Hier ist ein Beispiel für %, das als Platzhalterzeichen am Anfang eines Suchbegriffs verwendet wird:

  ```
  SELECT DATE, SUBJECT
    FROM DB2TX.SAMPLE
    WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"%name"') = 1
  ```

  Bei diesem Suchbegriff werden Textdokumente gefunden, die beispielsweise “Benutzername”, “Dateiname” und “Tabellenname” enthalten.

  % kann außerdem ein **ganzes Wort** repräsentieren. Im folgenden Beispiel werden Textdokumente gesucht, die Ausdrücke, wie z.&rbl;B. “graphic function” und “query function” enthalten.

  ```
  SELECT DATE, SUBJECT
    FROM DB2TX.SAMPLE
    WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"% function"') = 1
  ```

- _ repräsentiert **ein Zeichen** in einem Suchbegriff. Im folgenden Beispiel werden Textdokumente gesucht, die “CLOB” und “BLOB” enthalten.

  ```
  SELECT DATE, SUBJECT
         FROM DB2TX.SAMPLE
         WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"_LOB"') = 1
  ```

Nach Begriffen suchen, die bereits ein Platzhalterzeichen enthalten

Wenn Sie nach einem Begriff suchen wollen, der das Zeichen “%” oder das Zeichen “_” enthält, müssen Sie dem Zeichen ein sogenanntes *Escape-Zeichen* voranstellen und anschließend das Escape-Zeichen mit Hilfe des Schlüsselworts ESCAPE identifizieren.

Beispielsweise können Sie wie folgt nach dem Begriff “10% Zinsen” suchen:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        '"10!% Zinsen" ESCAPE "!"') = 1
```

Das Escape-Zeichen in diesem Beispiel ist `"!"`.

Suchargumente angeben

Nach Begriffen in beliebiger Reihenfolge suchen

Wenn Sie, wie im folgenden Beispiel gezeigt, nach “hard disk” suchen, finden Sie die beiden Begriffe nur, wenn sie benachbart und in der gezeigten Reihenfolge auftreten, unabhängig davon, welchen Indextyp Sie verwenden.

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"hard disk"') = 1
```

Um nach Begriffen in beliebiger Reihenfolge zu suchen, z.&rbl;B. wie in “data disks and hard drives”, verwenden Sie ein Komma, um die Begriffe voneinander zu trennen:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE, '("hard", "disk")') = 1
```

Nach Begriffen im selben Satz oder Absatz suchen

Hier ist ein Beispiel eines Suchargumente, mit dem Textdokumente gefunden werden, in denen die Suchbegriffe im selben Satz auftreten:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                 '"compress" IN SAME SENTENCE AS "decompress"') = 1
```

Sie können auch nach mehr als zwei Wörtern suchen, die zusammen auftreten. Im nächsten Beispiel wird nach mehreren Wörtern gesucht, die im selben Absatz auftreten:

```
SELECT DATE, SUBJECT
       FROM DB2TX.SAMPLE
       WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        '"compress" IN SAME PARAGRAPH AS "decompress"
                                              AND "encryption"') = 1
```

! ### Nach Begriffen in Abschnitten von strukturierten Dokumenten suchen

! Hier ist ein Beispiel für einen Suchbegriff, mit dem Textdokumente gefunden werden, in
! denen der Suchbegriff `Williams` im Unterabschnitt `author` des Abschnitts `play` von
! strukturierten Dokumenten auftritt. Die Dokumentstruktur ist durch das Modell `play`
! angegeben, das in einer Dokumentmodelldatei beschrieben ist. Weitere Informationen
! finden Sie in „Die Vorteile der Struktur eines Dokument nutzen“ auf Seite 68.

```
! SELECT DATE, SUBJECT
!        FROM DB2TX.SAMPLE
!        WHERE DB2TX.CONTAINS (COMMENTHANDLE,
!                        'MODEL play SECTIONS (play/author) "williams"') = 1
```

Nach Synonymen von Begriffen suchen

Bei einem linguistischen oder einem dualen Index können Sie Ihre Suchen flexibler gestalten, indem Sie nicht nur nach den angegebenen Suchbegriffen suchen, sondern auch nach Wörtern, die eine ähnliche Bedeutung haben. Wenn Sie beispielsweise nach dem englischen Wort “book” suchen, kann es auch nützlich sein, nach dessen Synonymen zu suchen.

Geben Sie dazu folgenden Befehl ein:

```
SELECT DATE, SUBJECT
       FROM DB2TX.SAMPLE
       WHERE DB2TX.CONTAINS (COMMENTHANDLE, 'SYNONYM FORM OF "book"') = 1
```

Wenn Sie SYNONYM FORM OF verwenden, wird davon ausgegangen, daß die Synonyme des Begriffs durch einen Operator für logische Oder verbunden sind, d.&rbl;h., das Suchargument wird folgendermaßen interpretiert:

```
"book" | "article" | "volume" | "manual"
```

Die Synonyme befinden sich in einem Wörterverzeichnis, das mit dem Text Extender geliefert wird. Der Standardwörterverzeichnis, das für Synonyme verwendet wird, ist immer US_ENGLISH, nicht die in den Textkonfigurationseinstellungen angegebene Sprache.

Sie können das Wörterverzeichnis für eine bestimmte Abfrage ändern, indem Sie eine andere Sprache angeben. Beispiel:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        'SYNONYM FORM OF GERMAN "programme"') = 1
```

Hinweis

Sie können das Schlüsselwort SYNONYM nicht verwenden, wenn im Suchbegriff Platzhalterzeichen auftreten oder wenn mit dem Suchargument NOT verwendet wird.

Eine linguistische Suche durchführen

Der Text Extender bietet die leistungsfähige Verarbeitung auf linguistischer Basis zum Durchführen einer Suche auf der Basis der von Ihnen gelieferten Suchbegriffe. Die linguistischen Funktionen werden angewendet, wenn der Index linguistisch ist oder wenn ein dualer Index mit dem Parameter STEMMED FORM OF verwendet wird. Die linguistischen Funktionen werden in Kapitel 3, „Verarbeitung auf linguistischer Basis“ auf Seite 19 beschrieben.

Ein Beispiel dafür ist die Suche nach einer Pluralform, z.&rbl;B. “utilities” und das Auffinden des Singulars, “utility”. Der Plural wird unter Verwendung eines englischen Wörterverzeichnisses auf seine Grundform `utility` reduziert, bevor die Suche beginnt. Das englische Wörterverzeichnis verfügt jedoch nicht über die Informationen zum Reduzieren von Variationen von Begriffen in anderen Sprachen auf ihre Grundform. Um nach dem Plural eines Begriffs in einer anderen Sprache zu suchen, müssen Sie das Wörterverzeichnis für diese Sprache verwenden. Wenn Sie beispielsweise GERMAN angeben, können Sie nach “geflogen” suchen und finden alle Variationen der Grundform “fliegen” - nicht nur “geflogen”, sondern auch “fliege”, “fliegt” usw.

Suchargumente angeben

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        'STEMMED FORM OF GERMAN "geflogen"') = 1
```

Hinweis

Beim Suchen in Dokumenten, die nicht in amerikanischem Englisch geschrieben sind, geben Sie die Sprache im Suchargument an, *unabhängig von der Standardsprache*.

Wenn Sie immer die Grundform eines Suchbegriffs angeben anstelle einer Variation davon, brauchen Sie keine Sprache anzugeben.

Um dies zu verstehen, bedenken Sie, was passiert, wenn der Text in Ihrer Datenbank indexiert wird. Wenn Sie einen linguistischen Index oder dualen Index verwenden, werden alle Variationen eines Begriffs auf ihre Grundform reduziert, bevor die Begriffe im Index gespeichert werden. Dies bedeutet, daß in der Tabelle DB2TX.SAMPLE, obwohl “decompress” im ersten Eintrag in der Spalte COMMENT steht und “decompression” im zweiten Eintrag steht, der Index nur die Grundform “decompress” enthält und diesen Begriff (oder dessen Variationen) als Inhalt von beiden Einträgen identifiziert.

Später wenn Sie nach der Grundform “decompress” suchen, finden Sie alle Variationen. Wenn Sie jedoch nach einer Variation, wie z.&rbl;B. “decompression” suchen, können Sie sie nicht direkt finden. Sie müssen ein entsprechendes Wörterverzeichnis für die Suche angeben, so daß die Variation zunächst in ihre Grundform umgesetzt werden kann.

Mit dem Booleschen Operator NOT suchen

Sie können den Booleschen Operator NOT verwenden, um bestimmte Textdokumente aus der Suche auszuschließen. Beispiel:

```
("compress", "compiler") & NOT "DB2"
```

Alle Textdokumente, die den Begriff “DB2” enthalten, sind von der Suche nach “compress” oder “compiler” ausgeschlossen.

Sie können den Operator NOT nicht in Kombination mit IN SAME SENTENCE AS oder IN SAME PARAGRAPH AS (in „Nach Begriffen im selben Satz oder Absatz suchen“ auf Seite 106 beschrieben) verwenden und auch nicht mit SYNONYM FORM OF (in „Nach Synonymen von Begriffen suchen“ auf Seite 106 beschrieben).

Sie können den Operator NOT nur mit einem Suchbegriff verwenden, d.&rbl;h., Sie können die Operatoren &, | und NOT nicht frei kombinieren (siehe „Suchargumentsyntax“ auf Seite 194).

Beispiel für die Verwendung von NOT, die **nicht** zulässig ist:

```
NOT("compress" & "compiler")
```

Zulässig ist:

```
NOT("compress" , "compiler")
```

Suche nach grober Übereinstimmung

Bei “Suchen nach grober Übereinstimmung” wird nach Wörtern gesucht, die ähnlich geschrieben werden wie der Suchbegriff. Diese Suche ist für Ngram-Indizes verfügbar.

Beispiel:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        'FUZZY FORM OF 2 "compress"') = 1
```

Diese Suche kann ein Vorkommen des Wortes mit Rechtschreibfehlern, `conpress`, finden.

Die Übereinstimmungsebene, in dem Beispiel “2”, gibt den Grad der Genauigkeit an. Fünf Ebenen werden unterstützt, wobei die Ebene 1 die niedrigste Übereinstimmung von etwa 20 Prozent angibt und Ebene 5 die strengste Übereinstimmung von etwa 90 Prozent angibt. Verwenden Sie eine Suche nach grober Übereinstimmung, wenn im Dokument möglicherweise Rechtschreibfehler vorkommen, was oft der Fall ist, wenn das Dokument mit Hilfe einer Einheit zur optischen Zeichenerkennung oder einer phonetischen Eingabe erstellt wurde.

Wortgrenzen respektieren

Die “BOUND”-Suche wurde für Koreanisch entwickelt. Sie stellt sicher, daß der Text Extender während der Suche Wortgrenzen respektiert. Beispiel:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        'BOUND "koreanischer_ausdruck"') = 1
```

Nach ähnlich klingenden Wörtern suchen

Bei “SOUND”-Suchen werden Wörter gefunden, die so ähnlich klingen wie das Suchargument. Dies ist nützlich, wenn Dokumente Wörter enthalten können, die ähnlich klingen, aber anders geschrieben werden. Der deutsche Name, der `mei-a` ausgesprochen wird, hat beispielweise mehrere Schreibweisen.

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                        'SOUNDS LIKE "Meyer"') = 1
```

Diese Suche findet Vorkommen von “Meyer”, “Mayer”, “Meier” und “Maier”.

Suchargumente angeben

Thesaurus-Suche

Die Thesaurus-Suche ist ein weiteres Beispiel für die leistungsfähigen Erweiterungsfunktionen für Suchbegriffe des Text Extender. Die zusätzlichen Begriffe, nach denen gesucht wird, werden aus einem Thesaurus genommen, den Sie selbst erstellen, so daß Sie die direkte Kontrolle darüber haben. Sie suchen beispielsweise nach “database” und könnten Begriffe wie “repository” und “DB2” finden.

Diese Art der Suche ist für bestimmte Bereiche konzipiert, in denen Sie häufig Suchen durchführen; Bereiche, für die sich die zeitliche Investition zum Erstellen eines Thesaurus lohnt, um bedeutend effektivere Suchergebnisse zu erzielen.

Weitere Informationen und eine Beschreibung zum Erstellen eines Thesaurus finden Sie in „Thesaurus-Konzepte“ auf Seite 32. Das Beispiel in Abb. 5 auf Seite 33 ist ein kleiner Auszug aus einem Thesaurus zum Thema Datenbanken. Er wird in den folgenden Beispielen verwendet, die die Syntax für die Verwendung der Thesaurus-Erweiterung demonstrieren.

Dieses Beispiel nimmt den Begriff “object relational database management system” und erweitert ihn, wobei alle *Exemplare* (Instances) dieses Begriffs, die im Thesaurus “myterms” gefunden werden, hinzugefügt werden. Hier wird “DB2” zur Suche hinzugefügt.

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
              'THESAURUS "myterms"
              EXPAND "INST"
              TERM OF "object relational database management system"
              ') = 1
```

Das nächste Beispiel nimmt den Begriff “document management system” und erweitert ihn, wobei alle *Synonyme* hinzugefügt werden. Ein Synonym wird gefunden, “library”.

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
              'THESAURUS "myterms"
              EXPAND "SYN"
              TERM OF "document management system"
              ') = 1
```

Freitext- und Hybridsuche

Die “Freitextsuche” ist eine Suche, in der der Suchbegriff als unformatierter Text ausgedrückt wird. Ein Ausdruck oder ein Satz beschreibt in natürlicher Sprache das Subjekt, nach dem gesucht werden soll. Die Reihenfolge von Wörtern in einer Freitextsuche ist nicht relevant. Darüber hinaus werden sogenannte *lexikalische Affinitäten* unterstützt. Bei der Abfrage sind dies bestimmte Wortpaare, die in einem Freitextabfragebegriff vorkommen und in der Dokumentensammlung vorkommen, wobei eine bestimmte Mindesthäufigkeit und ein bestimmter Mindestabstand gilt. Der Abstand für englische Dokumente ist fünf Wörter.

Beachten Sie, daß Platzhalterzeichen für Zeichen oder Wörter für Suchbegriffe in einem Freitextargument nicht unterstützt werden.

Beispiel:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
        'IS ABOUT "everything related to AIX installation"') = 1
```

Die Hybridsuche ist eine Kombination aus Boolescher Suche und Freitextsuche. Beispiel:

```
SELECT DATE, SUBJECT
  FROM DB2TX.SAMPLE
  WHERE DB2TX.CONTAINS (COMMENTHANDLE,
        '"DB2" & IS ABOUT "everything related to AIX installation"') = 1
```

Eine Feinanpassung für eine vorherige Suche durchführen

Wenn ein Suchargument zu viele Vorkommen findet, kann es oft nützlich sein, die Suche einzugrenzen oder eine *Feinanpassung* durchzuführen, wobei das ursprüngliche Suchargument mit einem zweiten Suchargument in einer Booleschen AND-Beziehung kombiniert wird.

Sie können eine Feinanpassung für Suchergebnisse durchführen, ohne die Funktion REFINE zu verwenden, indem Sie die Ergebnisse in einer Tabelle speichern und die nächste Suche für diese Tabelle ausführen. Ja nach der Anzahl der in Frage kommenden Begriffe ist diese Methode jedoch weniger effizient als das Speichern des letzten Suchargumentes und die Verwendung von REFINE.

Die folgenden Schritte zeigen, wie eine Suche durchgeführt wird und für diese anschließend unter Verwendung der Funktion REFINE eine Feinanpassung durchgeführt wird. Die Funktion REFINE gibt ein Suchargument zurück, das eine Boolesche AND-Kombination der zwei Eingabeparameter ist. Das kombinierte Suchargument, das von REFINE zurückgegeben wird, ist ein Wert vom Typ LONG VARCHAR.

Eine Feinanpassung für eine vorherige Suche durchführen

1. Erstellen Sie eine Tabelle für vorherige Suchargumente.

   Erstellen Sie eine Tabelle PREVIOUS_SEARCHES, in der die Suchargumente von bereits durchgeführten Suchen gespeichert werden sollen.

   ```
   CREATE TABLE PREVIOUS_SEARCHES (step INT,
                                   searchargument LONG VARCHAR)
   ```

   PREVIOUS_SEARCHES

STEP	SEARCHARGUMENT

2. Suchen Sie nach dem ersten Suchargument.

   Suchen Sie nach dem Wort “compress” in der Beispieltabelle.

   ```
   SELECT COMMENT
     FROM DB2TX.SAMPLE
     WHERE DB2TX.CONTAINS (COMMENTHANDLE, '"compress"') = 1
   ```

   Fügen Sie das Suchargument in die Tabelle PREVIOUS_SEARCHES ein, um es bei weiteren Schritten zu verwenden.

   ```
   INSERT INTO PREVIOUS_SEARCHES
     VALUES (1, '"compress"')
   ```

   PREVIOUS_SEARCHES

STEP	SEARCHARGUMENT
1	"compress"

3. Führen Sie eine Feinanpassung der Suche durch.

   Angenommen, die Suche gibt zu viele Textdokumente zurück, dann führen Sie eine Feinanpassung der Suche durch, indem Sie den vorherigen Suchbegriff unter Verwendung der Funktion REFINE mit dem Wort “compiler” kombinieren.

   ```
   WITH LAST_STEP(STEP_MAX)
        AS (SELECT MAX(STEP)
            FROM PREVIOUS_SEARCHES),
        LAST_SEARCH(LAST_SEARCH)
        AS (SELECT SEARCHARGUMENT
            FROM PREVIOUS_SEARCHES,LAST_STEP
            WHERE STEP = STEP_MAX)
   SELECT COMMENT
        FROM DB2TX.SAMPLE, LAST_SEARCH
     WHERE DB2TX.CONTAINS (COMMENTHANDLE,
                   DB2TX.REFINE(LAST_SEARCH, '"compiler"')) = 1
   ```

   Fügen Sie das angepaßte Suchargument in die Tabelle PREVIOUS_SEARCHES ein, um es bei weiteren Schritten zu verwenden.

```
INSERT INTO PREVIOUS_SEARCHES
  WITH LAST_STEP(STEP_MAX)
       AS (SELECT MAX(STEP)
           FROM PREVIOUS_SEARCHES)
  SELECT STEP_MAX+1, DB2TX.REFINE(SEARCHARGUMENT, '"compiler"')
    FROM PREVIOUS_SEARCHES, LAST_STEP
```

PREVIOUS_SEARCHES

STEP	SEARCHARGUMENT
1	"compress"
2	"compress" & "compiler"

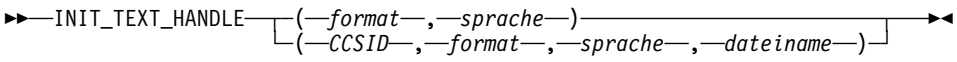
Sie können diesen Schritt wiederholen, bis die Anzahl der gefundenen Textdokumente klein genug ist.

Informationen in Kennungen setzen und extrahieren

Kennungen enthalten die CCSID (ID für codierten Zeichensatz), das Format und die Sprache ihrer Textdokumente. Kennungen für externe Dateien enthalten zusätzlich einen Zeiger auf die externe Datei. Diese Kennungen werden erstellt, wenn Sie eine Textspalte oder externe Dateien aktivieren.

Mit den in diesem Abschnitt beschriebenen UDFs können Sie die Textdaten in Kennungen setzen oder ändern.

Textdaten beim Einfügen von neuem Text setzen



Wenn Sie den Befehl ENABLE TEXT COLUMN ausführen, um eine Textspalte zu aktivieren, die bereits Text enthält, können Sie implizit das Format und die Sprache des Textes auf die Werte setzen, die in den Textkonfigurationseinstellungen angegeben sind. Diese Format- und Spracheinstellungen werden anschließend in der Kennung gespeichert. Wenn Sie andere Werte für Format und Sprache wünschen, können Sie sie explizit im Befehl ENABLE TEXT COLUMN angeben.

Wenn Sie den Befehl ENABLE TEXT FILES ausführen, können Sie auch die CCSID und den Standort des Dokuments festlegen.

Wenn Sie später eine Zeile mit Text einfügen, erstellt ein Einfügeauslöser eine Kennung und setzt das Textformat und die Sprache auf die Werte, die bei der Aktivierung der Textspalte verwendet wurden.

Informationen in Kennungen setzen und extrahieren

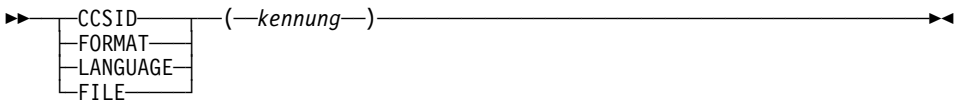
Um das Format und die Sprache auf andere Werte zu setzen, verwenden Sie die Funktion INIT_TEXT_HANDLE im Befehl INSERT. Während die Zeile eingefügt wird, erstellt die Funktion INIT_TEXT_HANDLE eine teilweise initialisierte Kennung, die die von Ihnen angegebenen Werte für Sprache und Format enthält. Der Einfügeauslöser füllt anschließend die anderen Werte in der Kennung aus.

Im folgenden Beispiel setzt die Funktion INIT_TEXT_HANDLE vorab die Sprache und das Format in einer initialisierten Kennung. Der Befehl INSERT stellt diese Kennung in die Spalte COMMENTHANDLE.

```
INSERT INTO DB2TX.SAMPLE (DOCID, COMMENT, COMMENTHANDLE)
       VALUES ('doc 101',
               'I have installed...',
               DB2TX.INIT_TEXT_HANDLE('AMI', 'GERMAN') )
```

Der Wert, der von INIT_TEXT_HANDLE zurückgegeben wird, ist ein Wert vom Typ DB2TEXTH oder DB2TEXTFH.

Informationen aus Kennungen extrahieren



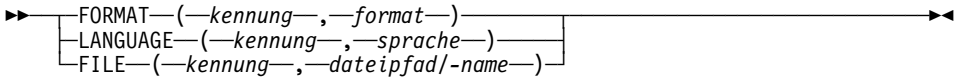
Hier ist ein Beispiel für die Extraktion einer CCSID aus einer Kennung:

```
SELECT DISTINCT DB2TX.CCSID(COMMENTHANDLE)
  FROM DB2TX.SAMPLE
```

Auf die gleiche Weise können Sie das Format oder die Sprache eines Textdokuments oder den Standort von externen Dateien extrahieren. Hier ist ein Beispiel für die Verwendung der Funktion FORMAT. Es gibt die Anzahl der ASCII-Dokumente (TDS-Dokumente) in der Beispieltabelle zurück.

```
SELECT COUNT(*)
  FROM DB2TX.SAMPLE
       WHERE DB2TX.FORMAT(COMMENTHANDLE) = 'TDS'
```

Informationen in Kennungen ändern



Mit den Funktionen FORMAT und LANGUAGE können auch die entsprechenden Angaben in einer Kennung geändert werden. Diese Funktionen geben die geänderte Kennung als einen Wert vom Typ DB2TEXTH, oder DB2TEXTFH zurück.

Das folgende Beispiel zeigt, wie die Spracheinstellung eines Textdokuments geändert wird.

```
UPDATE DB2TX.SAMPLE
       SET COMMENTHANDLE = DB2TX.LANGUAGE(COMMENTHANDLE, 'FRENCH')
       WHERE ...
```

Wenn Sie die UDF LANGUAGE erneut verwenden, können Sie die erfolgte Änderung sehen:

```
SELECT DISTINCT DB2TX.LANGUAGE(COMMENTHANDLE)
       FROM DB2TX.SAMPLE
```

Die Suchleistung verbessern

Die UDF SEARCH_RESULT nutzt das DB2-Konzept von Funktionen mit dem Wert Tabelle. Die UDF wird in der Klausel FROM einer SQL-Anweisung verwendet und gibt eine Zwischentabelle mit dem Suchergebnis des angegebenen Suchbegriffs zurück. Die Syntax des Suchbegriffs entspricht der Syntax, die in Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, beschrieben ist. Der Vorteil dieser UDF gegenüber CONTAINS oder RANK ist eine wesentliche Leistungsverbesserung, wenn große Tabellen beteiligt sind.

Die zurückgegebenen Tabelle hat die folgende Struktur:

```
! Spaltenname         Datentyp
! HANDLE              DB2TX.DB2TEXTH, DB2TX.DB2TEXTFH
! NUMBER_OF_MATCHES   INTEGER
! RANK                DOUBLE
```

Beispiel:

```
SET CURRENT FUNCTION PATH = db2tx
WITH REPHANDLE (MYDOCHANDLE) AS
  ( SELECT DB2TX.DB2TEXTH(PROTOTYPEHANDLE)
    FROM db2tx.textcolumns
    WHERE TABLESCHEMA = 'DB2TX'  AND
          TABLENAME   = 'SAMPLE' AND
          COLUMNNAME  = 'COMMENT'
  )
SELECT NUMBER_OF_MATCHES,RANK,HANDLE
   FROM REPHANDLE,
   TABLE(DB2TX.SEARCH_RESULT(MYDOCHANDLE,'"compress"')) T1
```

! `SELECT NUMBER_OF_MATCHES,RANK,HANDLE` führt dazu, daß alle drei Elemente zurückgegeben werden, Sie können sie aber in einer beliebigen Kombination angeben. Sie können beispielsweise `RANK` weglassen, um den hohen Verarbeitungsaufwand zu vermeiden, der damit verbunden ist.

Wenn Sie nur den Wert `HANDLE` benötigen, können Sie einfach `SELECT COUNT(*)` verwenden.

Die Suchleistung verbessern

Kapitel 6. Die API-Funktionen zum Suchen und Anzeigen verwenden

In diesem Kapitel wird beschrieben, wie die Such- und Anzeigefunktionen der Text Extender-API verwendet werden. Eine detaillierte Beschreibung dieser Funktionen finden Sie in Kapitel 11, „API-Funktionen zum Suchen und Anzeigen“ auf Seite 205. Beispiele für Programme, die die API-Funktionen verwenden, finden Sie in Kapitel 15, „Beispiel-API-Programm“ auf Seite 265. In demselben Kapitel wird eine Beispielanzeigefunktion, DesBrowseDocument, beschrieben.

Hinweis

Bevor Sie eine Suche durchführen, sollten Sie „Indextypen“ auf Seite 13 lesen. Eine Suche kann je nach verwendetem Indextyp zu unterschiedlichen Ergebnissen führen. Verwenden Sie den Befehl GET INDEX SETTINGS, um zu ermitteln, welche Textindextypen dem Text zugeordnet sind, in dem Sie suchen.

Sie sollten außerdem „Die Beispieltabelle DB2TX.SAMPLE“ auf Seite 96 lesen.

Ihre Anwendung einrichten

Ein Anwendungsprogramm, das die Text Extender-API verwendet, ist eine DB2 CLI-Anwendung, da für einige API-Funktionen als Eingabe eine Datenbankverbindungskennung erforderlich ist. Dies bedeutet, daß die Regeln, die für DB2 CLI-Anwendungen berücksichtigt werden müssen, auch für Anwendungen gelten, die die Text Extender-API verwenden.

Schließen Sie in Ihre Anwendung die Datei `des_ext.h` ein, die im Unterverzeichnis `include` des Text Extender-Installationsverzeichnisses geliefert wird.

Um Ihr Anwendungsprogramm mit der Text Extender-API zu verwenden, verbinden Sie Ihr Programm mit der API.

Eine UNIX-Anwendung verbinden

Sie müssen die Bibliothek `libdescl.a` mit Ihrer Anwendung verbinden. Diese Bibliothek befindet sich im Unterverzeichnis `lib` des Text Extender-Installationsverzeichnisses.

Eine OS/2- oder Windows-Anwendung verbinden

Alle Text Extender-API-Funktionen befinden sich in der Bibliothek für dynamisches Verbinden `desclapi.dll`. Die externen API-Funktionsaufrufe werden in der Importbibliothek `desclapi.lib` aufgelöst. Sie werden in den entsprechenden Unterverzeichnissen `dll` und `lib` des Text Extender-Installationsverzeichnisses geliefert.

Um Ihre Anwendung mit der Bibliothek für dynamisches Verbinden `desclapi.dll` der API zu verbinden, verwenden Sie `desclapi.lib` als Bibliotheksdatei.

Überblick über die API-Funktionen

Dies sind die Such- und Anzeigefunktionen; das erste ist eine Suchfunktion, die restlichen sind Anzeigefunktionen:

DesGetSearchResultTable
DesGetBrowseInfo
DesStartBrowseSession
DesOpenDocument
DesGetMatches
DesCloseDocument
DesEndBrowseSession
DesFreeBrowseInfo

Hinweis

Viele der API-Funktionen benötigen eine Verbindungskennung (hdbc). Sie müssen diese Kennung unter Verwendung der Funktion SQLConnect liefern, was allerdings nicht verhindert, daß Sie den Text Extender von Programmen in eingebettetem SQL aus aufrufen können. Im Handbuch *DB2 Call Level Interface Guide and Reference* wird beschrieben, wie CLI-Anweisungen und Anweisungen in eingebettetem SQL gemischt werden können.

Sie können die Such- und Anzeigefunktionen für folgende Aktionen verwenden:

Nach Text suchen

Text anzeigen

Nach Text suchen

In diesem Szenario wird nur die Suchfunktion DesGetSearchResultTable benötigt. Sie erhält als Eingabe ein Suchargument und einen Kennungsspaltennamen. Sie sucht nach Text und stellt Informationen zu den gefundenen Dokumenten in eine Ergebnistabelle, die Sie zuvor vorbereitet haben.

Diese Funktion wird in „Nach Text suchen“ auf Seite 120 näher beschrieben. Siehe auch Kapitel 15, „Beispiel-API-Programm“ auf Seite 265.

Text anzeigen

Verwenden Sie die folgenden Funktionen in der angegebenen Reihenfolge:

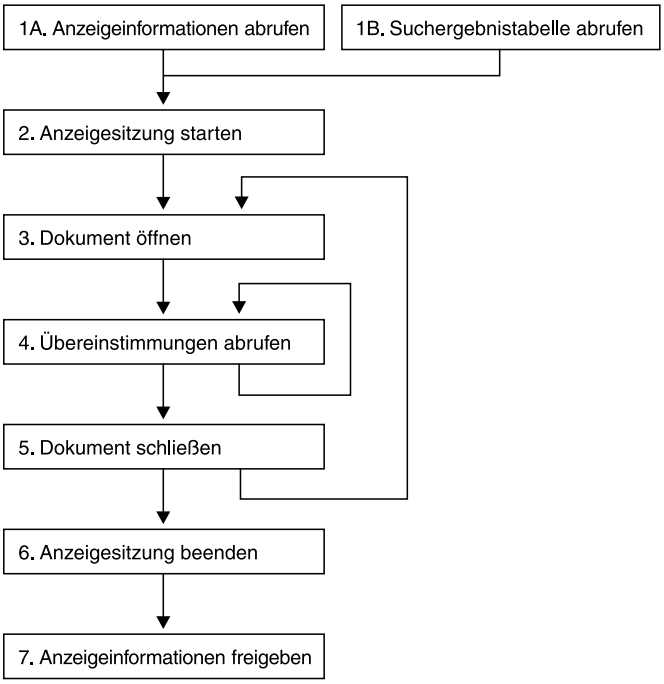


Abbildung 15. Reihenfolge für die Verwendung der API-Funktionen

Diese Funktionen erhalten Hervorhebungsinformationen, starten anschließend eine Anzeigesitzung, um ein Textdokument anzuzeigen, und heben die gefundenen Begriffe hervor.

In einer Anzeigesitzung können Sie weitere Dokumente öffnen und anzeigen, wobei dieselben Hervorhebungsinformationen verwendet werden. Diese Funktionen werden in „Text anzeigen“ auf Seite 121 näher beschrieben.

Nach Text suchen

Eine API-Funktion steht zur Verfügung, um nach Text zu suchen, die Funktion DesGetSearchResultTable.

Eine Suchergebnistabelle abrufen (DesGetSearchResultTable)

Die Funktion DesGetSearchResultTable empfängt ein Suchargument für die Suche in Textdokumenten in einer bestimmten Textspalte und speichert die Ergebnisse in einer Tabelle. Die Ergebnistabelle enthält die Kennung jedes gefundenen Dokuments. Sie enthält außerdem Informationen zur Rangordnung und die Anzahl der Übereinstimmungen, je nach angegebener Suchoption.

Sie können diese Informationen auch unter Verwendung der UDFs RANK und NO_OF_MATCHES abrufen. Beispiel:

```
INSERT INTO RESULT
       SELECT COMMENTHANDLE,
              RANK(COMMENTHANDLE, '"stored procedures"'),
              NO_OF_MATCHES(COMMENTHANDLE, '"stored procedures"')
       FROM DB2TX.SAMPLE
       WHERE  CONTAINS(COMMENTHANDLE, '"stored procedures"') = 1
```

Die Funktion DesGetSearchResultTable kann nur für Basistabellen verwendet werden, sie kann aber schneller sein als UDFs, wenn die Abfrage eine reine Textabfrage ist. Die Funktion geht direkt zum Text Extender-Server, um die Rangordnung und Anzahl von Übereinstimmungen abzurufen, und sie läuft nur in einer Schleife, um die Anzahl der gefundenen übereinstimmenden Dokumente zu ermitteln. Auf der anderen Seite im UDF-Beispiel wird die Funktion CONTAINS einmal für jede Zeile in der Tabelle aufgerufen. Anschließend werden für jede geeignete Zeile die Funktionen RANK und NO_OF_MATCHES aufgerufen. Für jedes gefundene Dokument werden drei separate Suchen durchgeführt.

Eingabe

Die Eingabeparameter sind folgende:

- Die Kennung für die Datenbankverbindung
- Die zu durchsuchende Tabelle
- Der Name der Kennungsspalte, die der zu durchsuchenden Textspalte zugeordnet ist
- Ein Suchargument
- Suchoptionen
- Eine Anzeigeoption (um Anzeigeinformationen zurückzugeben)
- Der Name der Tabelle, in der das Ergebnis gespeichert werden soll

Ausgabe

Wenn eine Anzeigeoption angegeben ist, gibt diese Funktion einen Zeiger auf Anzeigeinformationen zurück.

Text anzeigen

Diese Gruppe von Funktionen in Abb. 15 auf Seite 119 ermittelt, welche Begriffe hervorgehoben werden sollen. Sie startet anschließend eine Anzeigesitzung, öffnet ein Dokument und ruft Übereinstimmungsinformationen in Form eines Datenstroms ab, der von einem Anwendungsprogramm, das Ihren Browser aufruft, syntaktisch analysiert werden kann.

Anzeigeinformationen abrufen (DesGetBrowseInfo)

Die Funktion DesGetBrowseInfo empfängt ein Suchargument und eine Kennung. Sie gibt einen Zeiger auf Anzeigeinformationen zurück, die von DesStartBrowseSession benötigt werden. Zu den Anzeigeinformationen gehört eine Liste aller Begriffe, die hervorgehoben werden sollen.

Eine andere Methode zum Abrufen von Anzeigeinformationen besteht darin, die Anzeigeoption in der Funktion DesGetSearchResultTable anzugeben.

Eingabe

Die Eingabeparameter sind folgende:

Die Kennung für die Datenbankverbindung
Eine Kennung
Ein Suchargument

Ausgabe

Diese Funktion gibt einen Zeiger auf Anzeigeinformationen zurück.

Eine Anzeigesitzung starten (DesStartBrowseSession)

Die Funktion DesStartBrowseSession startet eine Anzeigesitzung, indem eine Umgebung hergestellt wird, die zum Anzeigen eines Textdokuments und zum Hervorheben von dessen Übereinstimmungen erforderlich ist. Sie empfängt einen Zeiger auf Anzeigeinformationen, entweder von DesGetBrowseInfo oder von DesGetSearchResultTable, und gibt eine Anzeigesitzungskennung für die Verwendung durch andere Anzeigefunktionen zurück.

Eingabe

Die Eingabeparameter sind folgende:

Ein Zeiger auf Anzeigeinformationen von DesGetBrowseInfo oder DesGetSearchResultTable
Eine Benutzer-ID
Ein Kennwort

Ausgabe

! Diese Funktion gibt eine Anzeigesitzungskennung zurück.

Text anzeigen

Ein Dokument öffnen (DesOpenDocument)

Die Funktion DesOpenDocument empfängt einen Anzeigesitzungszeiger, eine Kennung und die Option DES_FAST oder DES_EXTENDED, die die Art der Verarbeitung auf linguistischer Basis angibt, die für die Hervorhebung der gefundenen Begriffe verwendet werden soll. Siehe „Stufe 2: Erweiterte Übereinstimmung“ auf Seite 28. DES_FAST verwendet die Basistextanalyse, ohne die Verwendung eines Wörterverzeichnisses, um zu ermitteln, welche Begriffe hervorgehoben werden sollen. DES_EXTENDED verwendet die erweiterte Übereinstimmung. DesOpenDocument bereitet das Textdokument vor, das der Kennung entspricht, um den Dokumenttext und die Hervorhebungsinformationen abzurufen, und gibt eine Dokumentkennung zurück, die für den iterativen Aufruf von DesGetMatches verwendet wird.

Eingabe

Die Eingabeparameter sind folgende:

! Eine Anzeigesitzungskennung von DesStartBrowseSession
Eine Textkennung
Eine Übereinstimmungsoption: DES_FAST oder DES_EXTENDED

Ausgabe

Diese Funktion gibt eine Dokumentkennung zurück, die von DesGetMatches und DesCloseDocument verwendet wird.

Übereinstimmungen abrufen (DesGetMatches)

Die Funktion DesGetMatches gibt einen Zeiger auf Hervorhebungsinformationen für das Textdokument zurück, das durch eine Dokumentkennung beschrieben ist. Die Hervorhebungsinformationen sind ein Datenstrom. Dazu gehören der Textkontext (mindestens ein Absatz) und Informationen zum Hervorheben von Text in diesem Kontext. Der Datenstrom wird in „Datenstromsyntax“ auf Seite 213 beschrieben. Ein Anwendungsprogramm kann den Datenstrom syntaktisch analysieren und ihn unter Verwendung des eigenen Browsers des Benutzers verarbeiten. DesGetMatches gibt nur einen Abschnitt des Datenstroms zurück, wobei die Länge des Abschnitts in der Ausgabestruktur angegeben wird.

Eine Folge von Aufrufen von DesGetMatches ruft den Inhalt des gesamten Textdokuments ab. Wenn das Ende des Textdokuments erreicht ist, wird ein Anzeiger zurückgegeben.

Eingabe

Die Eingabeparameter sind folgende:

! Eine Anzeigesitzungskennung
Eine Dokumentkennung von DesOpenDocument

Ausgabe

Diese Funktion gibt einen Zeiger auf eine Struktur zurück, die den Datenstromabschnitt und dessen Länge enthält.

Ein Dokument schließen (DesCloseDocument)

Die Funktion DesCloseDocument schließt ein Textdokument, das von DesOpenDocument geöffnet wurde, und gibt den Speicher frei, der während der Rückgabe des Dokumenttextes und der Hervorhebungsinformationen zugeordnet wurde.

Eingabe

Die Eingabeparameter sind folgende:

! Eine Anzeigesitzungskennung

Eine Dokumentkennung von DesOpenDocument

Ausgabe

! Keine

Eine Anzeigesitzung beenden (DesEndBrowseSession)

Die Funktion DesEndBrowseSession beendet eine Anzeigesitzung, die von DesStartBrowseSession gestartet wurde, und gibt den Speicher frei, der für die Anzeigesitzung zugeordnet wurde.

Eingabe

Der Eingabeparameter ist folgender:

! Eine Anzeigesitzungskennung

Ausgabe

! Keine

Die Anzeigeinformationen freigeben (DesFreeBrowseInfo)

Die Funktion DesFreeBrowseInfo gibt den Speicher frei, der für die Anzeigeinformationen durch DesGetBrowseInfo zugeordnet wurde.

Eingabe

Der Eingabeparameter ist folgender:

Ein Zeiger auf die Anzeigeinformationen

Ausgabe

! Keine

Text anzeigen

Teil 2. Referenz

Kapitel 7. Verwaltungsbefehle für den Client

Befehl	Zweck	Seite
?	Hilfe für den Befehlszeilenprozessor	128
CHANGE INDEX SETTINGS	Ändert die Merkmale eines Indexes	129
CHANGE TEXT CONFIGURATION	Ändert die Textkonfigurationseinstellungen	130
CONNECT	Stellt eine Verbindung zu einer Datenbank her	133
DELETE INDEX EVENTS	Löscht Indexereignisse aus einer Protokolltabelle	134
DISABLE DATABASE	Inaktiviert eine Datenbank für die Verwendung durch den Text Extender	135
DISABLE TEXT COLUMN	Inaktiviert eine Textspalte für die Verwendung durch den Text Extender und löscht den zugeordneten Index	136
DISABLE TEXT FILES	Inaktiviert Textdateien für die Verwendung durch den Text Extender und löscht den zugeordneten Index	137
DISABLE TEXT TABLE	Inaktiviert eine Tabelle für die Verwendung durch den Text Extender und löscht die Indizes, die der Tabelle zugeordnet sind	138
ENABLE DATABASE	Bereitet eine Datenbank für die Verwendung durch den Text Extender vor	139
ENABLE TEXT COLUMN	Bereitet eine Textspalte für die Verwendung durch den Text Extender vor und erstellt einen individuellen Textindex für die Spalte	139
ENABLE TEXT FILES	Bereitet Textdateien für die Verwendung durch den Text Extender vor und erstellt einen individuellen Textindex für die Dateien	147
ENABLE TEXT TABLE	Erstellt einen gemeinsamen Textindex für eine Tabelle	150
GET ENVIRONMENT	Zeigt die aktuellen Einstellungen der Umgebungsvariablen an	154
GET INDEX SETTINGS	Zeigt die Merkmale eines Indexes an	154
GET INDEX STATUS	Zeigt Statusinformationen für einen Index an	156
GET STATUS	Zeigt den aktivierten Status von Datenbanken, Tabellen und Spalten an	157
GET TEXT CONFIGURATION	Zeigt die Textkonfigurationseinstellungen an	157
GET TEXT INFO	Zeigt die Textdaten für eine Textspalte an	158
QUIT	Verläßt den Modus für den Verwaltungsbefehlszeilenprozessor	158
! REORGANIZE INDEX	Reorganisiert einen Index, um die Sucheffizienz zu verbessern	159
RESET INDEX STATUS	Setzt den Status eines Indexes zurück, um zu ermöglichen, daß er erneut verwendet wird	160
UPDATE INDEX	Aktualisiert einen Textindex	160

In diesem Kapitel wird die Syntax von Verwaltungsbefehlen für den Client beschrieben. Die Client-Verwaltung besteht aus Tasks, die Sie ausführen müssen, bevor Sie mit der Suche in Textdokumenten beginnen können, und aus Verwaltungs-Tasks. Kapitel 4, „Verwaltung“ auf Seite 43 beschreibt die Verwendung dieser Befehle.

Bevor Sie diese Befehle verwenden, starten Sie den Text Extender-Befehlszeilenprozessor, indem Sie den Befehl `db2tx` eingeben. Dieser Befehl aktiviert einen interaktiven Eingabemodus, in dem alle nachfolgenden Befehle als Text Extender-Befehle interpretiert werden. Normalerweise starten Sie den Befehlsprozessor zur gleichen Zeit wie DB2.

Um diesen Modus zu verlassen, geben Sie QUIT ein.

Wie bei DB2 können Sie Befehle direkt von der Eingabeaufforderung des Betriebssystems aus eingeben, indem Sie das Präfix `db2tx` verwenden. Außerdem können Sie Anforderungen für den Befehlszeilenprozessor in eine Befehlsdatei für eine Shell-Prozedur einschließen.

Hinweis

Ein Befehl mit dem Präfix `db2tx` führt dazu, daß eine Verbindung zur Standarddatenbank hergestellt wird, die in der Umgebungsvariablen DB2DBDFT angegeben ist. Die folgende Befehlsfolge aktiviert nicht die Datenbank MYDATABASE, sondern statt dessen die Standarddatenbank.

```
db2tx "CONNECT TO MYDATABASE"
db2tx "ENABLE DATABASE"
```

Hilfe für den Befehlszeilenprozessor

Um eine Liste der Verwaltungsbefehle aufzurufen, geben Sie folgendes ein:

```
db2tx "?"
```

Um die Syntax zu einem individuellen Befehl aufzurufen, geben Sie folgendes ein:

`db2tx "? befehl"`

Beispiel:

```
db2tx "? CHANGE INDEX SETTINGS"
```

CHANGE INDEX SETTINGS

Dieser Befehl ändert die Merkmale eines Indexes, **nachdem** die Datenbank aktiviert wurde.

Um die Standardeinstellungen zu ändern, die verwendet werden, wenn eine Datenbank das erste Mal aktiviert wird, verwenden Sie „CHANGE TEXT CONFIGURATION“ auf Seite 130.

Berechtigung

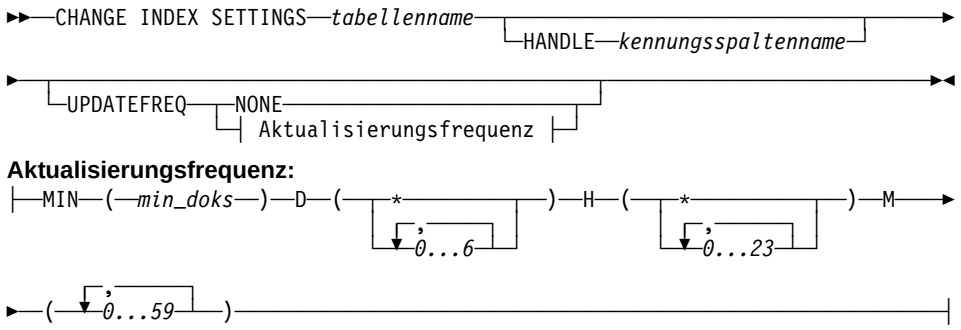
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalte enthält, deren Indexaktualisierungsfrequenz geändert werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Indexaktualisierungsfrequenz geändert werden soll. Dies ist erforderlich, wenn die Textspalte einen eigenen Index hat, d.&rbl;h., wenn der Index unter Verwendung des Befehls ENABLE TEXT COLUMN erstellt wurde.

CHANGE TEXT CONFIGURATION (Befehl)

Wenn jedoch der Index unter Verwendung des Befehls ENABLE TEXT TABLE erstellt wurde, d.&rbl;h., die Tabelle hat einen Textindex für alle Textspalten, wird dieses Schlüsselwort ignoriert.

UPDATEFREQ Aktualisierungsfrequenz Die Indexaktualisierungsfrequenz in bezug auf, wann die Aktualisierung durchgeführt werden soll und wie viele Textdokumente mindestens in der Protokolltabelle stehen müssen. Wenn zum bestimmten Zeitpunkt nicht genügend Textdokumente in der Protokolltabelle stehen, wird der Index nicht aktualisiert.

Die Syntax wird in „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262 beschrieben.

NONE Keine weiteren Indexaktualisierungen werden durchgeführt. Dies ist für eine Textspalte gedacht, an der keine weiteren Änderungen vorgenommen werden.

Wenn Sie das Schlüsselwort UPDATEFREQ nicht angeben, bleiben die Einstellungen für die Frequenz unverändert.

CHANGE TEXT CONFIGURATION

Dieser Befehl ändert die Standardeinstellungen der Textkonfiguration, die verwendet wird, wenn die Datenbank aktiviert wird. Es handelt sich um die Einstellungen zur *Textkonfiguration*. Die anfänglichen Textkonfigurationseinstellungen bei der Installation des Text Extender werden in „Textkonfigurationseinstellungen“ auf Seite 254 beschrieben.

Um diese Einstellungen für eine bestimmte Datenbank zu ändern, **nachdem** die Datenbank aktiviert wurde, verwenden Sie „CHANGE INDEX SETTINGS“ auf Seite 129.

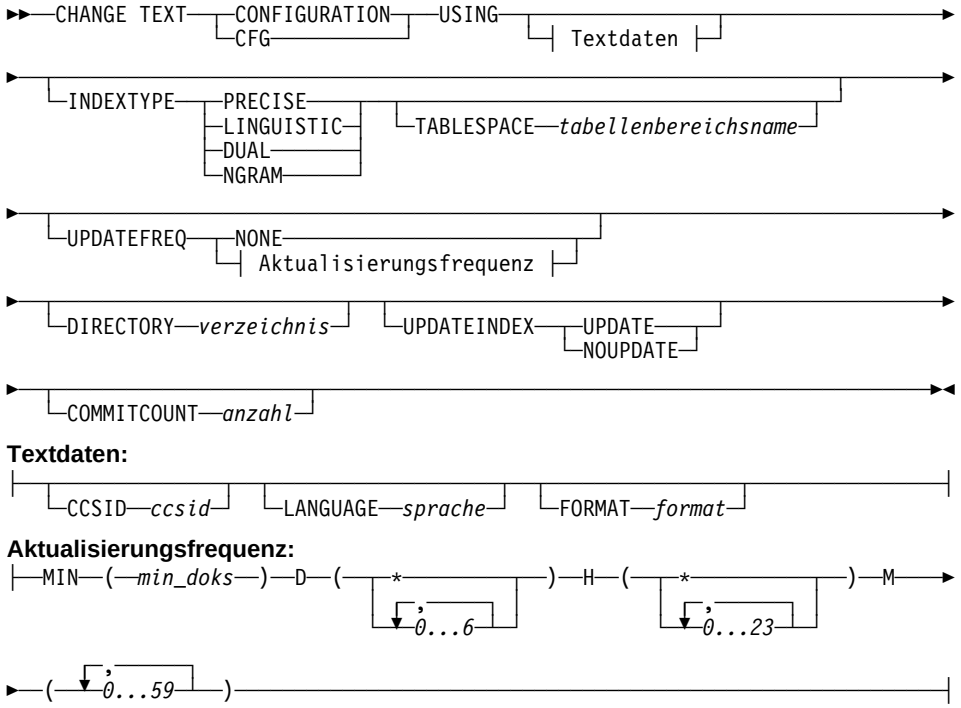
Berechtigung

Sie müssen über folgende Berechtigung verfügen:

SYSADM-Berechtigung

DBADM-Berechtigung

Befehlssyntax



Befehlsparameter

INDEXTYPE Um den Standardindextyp zu ändern, eine der folgenden Optionen auswählen. Weitere Informationen finden Sie in „Indextypen“ auf Seite 13.

PRECISE Die Begriffe werden exakt so indexiert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC Die Begriffe werden linguistisch verarbeitet, bevor sie indexiert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

DUAL Begriffe werden exakt so indexiert, wie sie in den Textdokumenten vorkommen und sie werden außerdem indexiert, nachdem sie linguistisch verarbeitet wurden. Bei der Suche können Sie für jeden Begriff entscheiden, ob Sie nach dem präzisen Begriff oder nach dem linguistisch verarbeiteten Begriff suchen wollen.

NGRAM Die Begriffe werden indexiert, indem Zeichengruppen syntaktisch analysiert werden, statt ein Wörterverzeichnis zu verwenden. Dieser Wörterverzeichnistyp ist verbindlich, wenn die Dokumente, die Sie indexieren, DBCS-Zeichen enthalten. Ein Ngram-Index kann aber auch für SBCS-Dokumente verwendet werden.

TABLESPACE tabellenbereichsname Geben Sie den Namen eines bestehenden Tabellenbereichs an. Der Tabellenbereich wird verwendet, um indexspezifische Verwaltungstabellen zu halten, die vom Text Extender (wie die Protokolltabellen) erstellt wurden. Verwenden Sie für große Tabellen einen separaten Tabellenbereich. Wenn Sie keinen Tabellenbereich angeben, werden die Tabellen im DB2-Standardtabellenbereich erstellt.

UPDATEFREQ Aktualisierungsfrequenz Die Indexaktualisierungsfrequenz in bezug auf, wann die Aktualisierung durchgeführt werden soll und wie viele Textdokumente mindestens in der Protokolltabelle stehen müssen. Wenn zum bestimmten Zeitpunkt nicht genügend Textdokumente in der Protokolltabelle stehen, wird der Index nicht aktualisiert.

Die Syntax wird in „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262 beschrieben.

NONE Keine weiteren Indexaktualisierungen werden durchgeführt. Dies ist für eine Textspalte gedacht, an der keine weiteren Änderungen vorgenommen werden.

DIRECTORY verzeichnis Das Verzeichnis, in dem der Textindex gespeichert werden soll.

UPDATEINDEX Ein Schlüsselwort, das festlegt, ob die Textdokumente sofort nach Abschluß des Befehls, der diese Option verwendet, indexiert werden sollen, ohne auf die nächste regelmäßige Indexierung zu warten, die durch UPDATEFREQ gesetzt ist. Die Befehle, die diese Option verwenden, sind ENABLE TEXT COLUMN und ENABLE TEXT FILES.

UPDATE Das Indexieren der Textdokumente findet sofort statt, nachdem der Befehl abgeschlossen ist.

NOUPDATE Das Indexieren findet zu dem Zeitpunkt statt, der durch die Einstellungen der Aktualisierungsfrequenz entweder in diesem Befehl durch UPDATEFREQ oder durch die Textkonfigurationseinstellung festgelegt ist.

COMMITCOUNT anzahl Ein Wert zwischen 500 und 1000000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen eine sofortige DB2-Commit-Anweisung ausgegeben wird. Dadurch können Situationen vermieden werden, in denen nicht genügend Protokollspeicherbereich vorhanden ist, wenn große Tabellen oder Spalten bzw. eine große Anzahl an externen Dateien aktiviert werden.

CCSID ccsid Die ID für den codierten Zeichensatz (CCSID), die beim Indexieren von Textdokumenten verwendet werden soll.

Informationen zu den CCSIDs, die unterstützt werden können, finden Sie in „CCSIDs“ auf Seite 259.

LANGUAGE sprache Die Sprache, in der der Text geschrieben ist. Hierdurch wird festgelegt, welches Wörterverzeichnis beim Indexieren von Textdokumenten und beim Suchen in Textdokumenten verwendet wird. Kapitel 3, „Verarbeitung auf linguistischer Basis“ auf Seite 19 beschreibt, wie Wörterverzeichnisse verwendet werden.

Die unterstützten Sprachen werden in „Sprachen“ auf Seite 258 aufgelistet.

FORMAT format Der Typ des gespeicherten Textdokuments, z.&rbl;B. WordPerfect oder ASCII. Der Text Extender benötigt die Informationen beim Indexieren von Dokumenten. Die unterstützten Dokumentformate werden in „Formate“ auf Seite 256 aufgelistet.

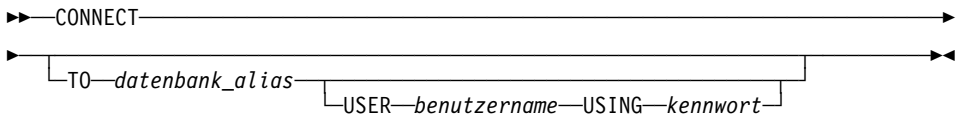
CONNECT

Dieser Befehl verbindet den Text Extender mit einer Datenbank.

Berechtigung

CONNECT-Berechtigung für die Datenbank.

Befehlssyntax



Befehlsparameter

TO datenbank_alias Die Datenbank, zu der eine Verbindung hergestellt werden soll. Der Name muß ein Aliasname für die Datenbank sein und kein Datenbankname.

USER benutzername Wenn kein Benutzername angegeben ist, wird er vom Betriebssystem abgerufen.

USING kennwort Ein Kennwort ist nur erforderlich, wenn ein Benutzername angegeben ist.

DELETE INDEX EVENTS (Befehl)

CONNECT kein operand Wenn Sie keinen Operanden angeben und keine Verbindung zu einer Datenbank besteht, stellt CONNECT eine implizite Verbindung zu der Datenbank her, die in der Umgebungsvariablen DB2DBDFT angegeben ist. Wenn Sie keinen Operanden angeben und eine Verbindung zu einer Datenbank besteht, zeigt CONNECT Informationen zur aktuellen Datenbank an.

Hinweise zur Verwendung

Wenn Sie einen Befehl DB2TX ausgeben, ohne bereits mit einer Datenbank verbunden zu sein, stellt der Text Extender eine Verbindung zur Datenbank her, die in der Umgebungsvariablen DB2DBDFT angegeben ist.

Um explizit eine Verbindung zu einer bestimmten Datenbank herzustellen, geben Sie den Befehl CONNECT TO ein.

Sie können nur jeweils mit einer Datenbank verbunden sein, diese Datenbank wird als aktuelle Datenbank bezeichnet. Im interaktiven Modus bleibt eine Verbindung bestehen, bis die Datenbank durch eine andere Anweisung CONNECT TO gewechselt wird oder bis ein Befehl QUIT ausgegeben wird. Im Befehlszeilenmodus hat der Befehl CONNECT keine Auswirkungen.

Wenn Sie einen OS/2-Client verwenden und zum ersten Mal eine Verbindung zu einer Datenbank herstellen, nachdem Sie das System gestartet haben, werden Sie von der Benutzerprofilverwaltung (UPM) aufgefordert, eine Benutzer-ID und ein Kennwort einzugeben. Wenn Sie bereits angemeldet sind, übernimmt die UPM die Benutzer-ID und das Kennwort aus dieser Sitzung und versucht, eine Verbindung zum fernen Server herzustellen. Siehe „Den Text Extender-Befehlszeilenprozessor starten (wahlfrei)“ auf Seite 48.

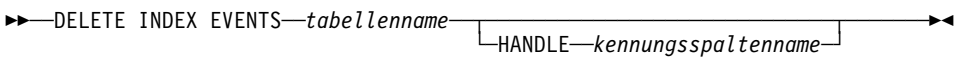
DELETE INDEX EVENTS

Dieser Befehl löscht Indexierungsereignisse aus der Protokolltabelle eines Indexes für eine bestimmte Kennungsspalte oder Tabelle.

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, deren Fehlerereignisse aus der Protokolltabelle gelöscht werden sollen. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Nachrichten aus der Protokolltabelle gelöscht werden sollen.

Hinweise zur Verwendung

Wenn der Name einer Kennungsspalte angegeben ist, werden nur die Indexierungsereignisse für diese Spalte gelöscht.

DISABLE DATABASE

Dieser Befehl setzt alle Vorbereitungsarbeiten, die der Text Extender für eine Datenbank durchgeführt hat, zurück und inaktiviert alle Texttabellen für die Verwendung durch den Text Extender.

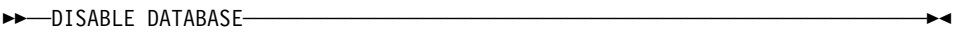
Berechtigung

Sie müssen über folgende Berechtigung verfügen:

SYSADM-Berechtigung

DBADM-Berechtigung

Befehlssyntax



Befehlsparameter

Keine

Hinweise zur Verwendung

Dieser Befehl setzt die verbundene Datenbank zurück, so daß sie nicht länger vom Text Extender durchsucht werden kann, d.&rbl;h., alle Text Extender-Texttabellen und -Textspalten in der Datenbank werden inaktiviert. Alle Änderungen, die in der Datenbank durchgeführt wurden, um Text Extender-Texttabellen, -Textspalten und externe Dateien zu aktivieren, werden zurückgesetzt, alle zugehörigen Textindizes werden gelöscht, die Text Extender-Katalogsicht TEXTCOLUMNS in der Datenbank wird gelöscht und alle Text Extender-Auslöser werden gelöscht.

DISABLE TEXT COLUMN (Befehl)

DISABLE TEXT COLUMN

Dieser Befehl inaktiviert eine Textspalte für die Verwendung durch den Text Extender.

Berechtigung

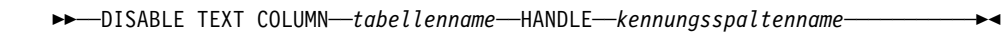
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die zu inaktivierende Spalte enthält. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, die für die Verwendung durch den Text Extender inaktiviert werden soll.

Hinweise zur Verwendung

Der Index wird gelöscht.

Die Protokolltabelle, die zum Aufzeichnen von Änderungen (Einfügungen, Aktualisierungen und Löschungen) in der Kennungsspalte verwendet wurde, wird gelöscht.

Die Auslöser, die Einträge in die Protokolltabelle schreiben, werden gelöscht.

Die Kennungsspalte wird nicht geändert.

DISABLE TEXT FILES

Dieser Befehl inaktiviert eine Gruppe von externen Textdateien für die Verwendung durch den Text Extender.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Kennungsspalte für die externen Textdateien enthält, die inaktiviert werden sollen. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, die für die Verwendung durch den Text Extender inaktiviert werden soll.

Hinweise zur Verwendung

Der Index wird gelöscht.

Die Protokolltabelle, die zum Aufzeichnen von Änderungen (Einfügungen, Aktualisierungen und Löschungen) in der Kennungsspalte verwendet wurde, wird gelöscht. Die Auslöser, die Einträge in die Protokolltabelle schreiben, werden ebenfalls gelöscht.

DISABLE TEXT TABLE (Befehl)

DISABLE TEXT TABLE

Dieser Befehl inaktiviert alle Textspalten in einer Tabelle für die Verwendung durch den Text Extender.

Berechtigung

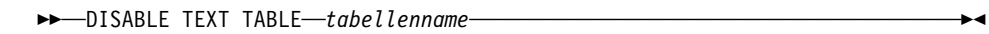
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die zu inaktivierende Spalte enthält. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

Hinweise zur Verwendung

Dieser Befehl inaktiviert alle Textspalten in der Tabelle für die Verwendung durch den Text Extender.

Wenn die Textspalten in dieser Tabelle einzeln durch ENABLE TEXT COLUMN aktiviert wurden, löscht der Befehl alle zugehörigen Textindizes. (Um die Textspalten einzeln zu inaktivieren und ihre zugehörigen Textindizes einzeln zu löschen, verwenden Sie den Befehl DISABLE TEXT COLUMN.) Wenn die Textspalten in dieser Tabelle zusammen mit dem Befehl ENABLE TEXT TABLE aktiviert wurden, existiert ein gemeinsamer Index für alle diese Textspalten. Dieser Befehl löscht den gemeinsamen Index.

Die Protokolltabellen, die zum Aufzeichnen von Änderungen (Einfügungen, Aktualisierungen und Löschungen) in der Kennungsspalte verwendet wurden, werden gelöscht. Die Auslöser, die Einträge in die Protokolltabelle schreiben, werden gelöscht.

ENABLE DATABASE

! Der Befehl aktiviert die aktuelle Datenbank zum Speichern von Textdaten.

Berechtigung

Sie müssen über folgende Berechtigung verfügen:

SYSADM-Berechtigung

DBADM-Berechtigung

Befehlssyntax

Befehlsparameter

Keine

Hinweise zur Verwendung

Sie müssen mit einer Datenbank verbunden sein, explizit oder implizit, bevor Sie diesen Befehl ausgeben können (siehe „CONNECT“ auf Seite 133). Dieser Befehl bereitet die verbundene Datenbank für die Verwendung durch den Text Extender vor. Dieser Schritt muß ausgeführt werden, bevor eine Text Extender-Texttabelle oder -Textspalte in der Datenbank aktiviert werden kann.

ENABLE DATABASE erstellt die Text Extender-Katalogsicht DB2TX.TEXTINDEXES, die in „Mit der Text Extender-Katalogsicht arbeiten“ auf Seite 89 beschrieben wird, und die Katalogsicht DB2TX.TEXTCOLUMNS, die für “Leistungsabfragen” verwendet wird.

Der Befehl erstellt außerdem Textkonfigurationseinstellungen, die in „Textkonfigurationseinstellungen“ auf Seite 254 beschrieben werden.

Einige andere Verwaltungsaufgaben werden ausgeführt, wie z.&rbl;B. die Deklaration von UDTs und UDFs.

ENABLE TEXT COLUMN

Dieser Befehl aktiviert eine Textspalte für die Verwendung durch den Text Extender.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

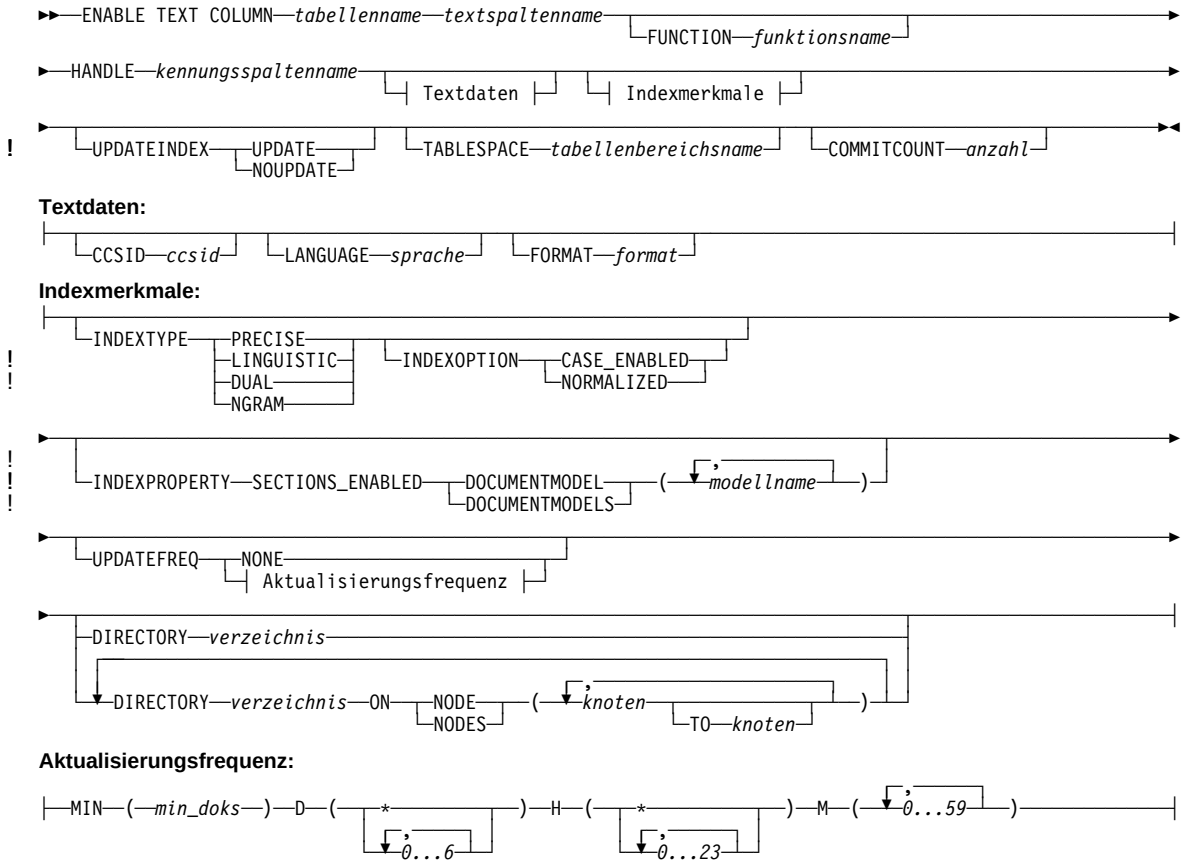
ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

ENABLE TEXT COLUMN (Befehl)

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die zu aktivierende Spalte enthält. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

textspaltenname Der Name der Spalte, die für die Verwendung durch den Text Extender aktiviert werden soll. Diese Spalte muß vom Typ CHAR, VARCHAR, LONG VARCHAR, CLOB, DBCLOB, GRAPHIC, VARGRAPHIC oder LONG VARGRAPHIC sein. Andernfalls verwenden Sie FUNCTION, um den Dokumenttyp umzusetzen.

FUNCTION funktionsname Das Name einer benutzerdefinierten Funktion, die von den Text Extender-Bibliotheksfunktionen für den Zugriff auf Textdokumente verwendet werden soll, die in einer Spalte stehen, die nicht vom Typ CHAR, VARCHAR, LONG VARCHAR, CLOB, DBCLOB, GRAPHIC, VARGRAPHIC oder LONG VARGRAPHIC ist. Weitere Informationen finden Sie in „Textspalten für einen nicht unterstützten Datentyp aktivieren“ auf Seite 66.

HANDLE kennungsspaltenname Der Name einer Kennungsspalte, die für die Verwendung durch die UDFs des Text Extender zur Tabelle hinzugefügt werden soll.

CCSID ccsid Die ID für den codierten Zeichensatz (CCSID), die beim Indexieren von Textdokumenten verwendet werden soll.

Wenn dieses Schlüsselwort nicht angegeben ist, wird die CCSID verwendet, die in den Textkonfigurationseinstellungen angegeben ist. Nachfolgende Änderungen an den Textkonfigurationseinstellungen werden ignoriert. Der verwendete Wert ist der Wert, der bei der Aktivierung der Spalte existierte, nicht der Wert, der beim Indexieren von Textdokumenten existierte.

Informationen zu anderen CCSIDs, die unterstützt werden können, finden Sie in „CCSIDs“ auf Seite 259.

LANGUAGE sprache Die Sprache, in der der Text geschrieben ist. Hierdurch wird festgelegt, welches Wörterverzeichnis beim Indexieren von Textdokumenten und beim Suchen in Textdokumenten verwendet wird. Kapitel 3, „Verarbeitung auf linguistischer Basis“ auf Seite 19, beschreibt, wie Wörterverzeichnisse verwendet werden.

Dieses Schlüsselwort gibt einmal die Sprache für die gesamte Spalte an. Sie können diesen Wert für einzeln eingefügte Textdokumente überschreiben, indem Sie die Funktion INIT_TEXT_HANDLE in einer Anweisung INSERT verwenden.

Wenn dieses Schlüsselwort nicht angegeben ist, wird die Sprache verwendet, die in den Textkonfigurationseinstellungen angegeben ist. Nachfolgende Änderungen an den Textkonfigurationseinstellungen werden ignoriert. Der verwendete Wert ist der Wert, der bei der Aktivierung der Spalte existierte, nicht der Wert, der beim Indexieren von Textdokumenten existierte.

Die unterstützten Sprachen werden in „Sprachen“ auf Seite 258 aufgelistet.

ENABLE TEXT COLUMN (Befehl)

FORMAT format Der Typ des gespeicherten Textdokuments, z.&rbl;B. WordPerfect oder ASCII. Der Text Extender benötigt die Informationen beim Indexieren von Dokumenten. Die unterstützten Dokumentformate werden in „Formate“ auf Seite 256 aufgelistet.

! Folgende Dokumentformate werden für strukturierte Dokumente unterstützt:

! • ASCII_SECTIONS

! Dokumente im Format ASCII_SECTIONS können keine
! verschachtelten Abschnitte enthalten. (Informationen zu
! verschachtelten Abschnitten finden Sie in „Die Vorteile der Struktur
! eines Dokument nutzen“ auf Seite 68.) Ein Startbefehl für einen
! Abschnitt wird durch den nächsten Startbefehl beendet.

! • HTML

! Ein Beispiel für eine Dokumentmodelldatei wird für HTML-Dokumente
! geliefert. Es enthält eine Untergruppe der Standard-HTML-Definitionen,
! die Sie gegebenenfalls ändern können. HTML-Dokumente können
! keine verschachtelten Abschnitte enthalten.

! • XML

! Zur Verarbeitung von XML-Dokumenten gehört die Bewertung der
! Dokumenttypendefinition (DTD). Das dem Dokument zugeordnete
! Modell wird der DTD gegenübergestellt. Wenn die Befehle, die in der
! Dokumentmodelldatei definiert sind, nicht in der DTD definiert sind,
! wird das Dokument nicht indexiert. Wenn kein Modell für eine erkannte
! DTD definiert wurde, wird das Dokument nicht indexiert.
! XML-Dokumente können verschachtelten Abschnitte enthalten.

! Für diese Formate müssen Sie die Strukturinformationen in einer
! Dokumentmodelldatei angeben. Siehe „Die Vorteile der Struktur eines
! Dokument nutzen“ auf Seite 68. Wenn das Format TDS und
! INDEXPROPERTY SECTION_ENABLED angegeben sind, wird davon aus-
! gegangen, daß das Dokumentformat ASCII_SECTIONS ist.

Befehle, die nicht in der Dokumentmodelldatei definiert sind, werden auf normale Art und Weise gemäß dem Indextyp indexiert.

Dieses Schlüsselwort gibt einmal das Format für die gesamte Spalte an. Sie können diesen Wert für einzeln eingefügte Textdokumente überschreiben, indem Sie die Funktion INIT_TEXT_HANDLE in einer Anweisung INSERT verwenden.

Wenn dieses Schlüsselwort nicht angegeben ist, wird das Format verwendet, das in den Textkonfigurationseinstellungen angegeben ist. Nachfolgende Änderungen an den Textkonfigurationseinstellungen werden ignoriert. Der verwendete Wert ist der Wert, der bei der Aktivierung der Spalte existierte, nicht der Wert, der beim Indexieren von Textdokumenten existierte.

INDEXTYPE Der Typ des zu erstellenden Indexes. Weitere Informationen finden Sie in „Indextypen“ auf Seite 13.

PRECISE Die Begriffe werden exakt so indexiert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC Die Begriffe werden linguistisch verarbeitet, bevor sie indexiert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

DUAL Begriffe werden exakt so indexiert, wie sie in den Textdokumenten vorkommen und sie werden außerdem indexiert, nachdem sie linguistisch verarbeitet wurden. Bei der Suche können Sie für jeden Begriff entscheiden, ob Sie nach dem präzisen Begriff oder nach dem linguistisch verarbeiteten Begriff suchen wollen.

NGRAM Die Begriffe werden indexiert, indem Zeichengruppen syntaktisch analysiert werden, statt ein Wörterverzeichnis zu verwenden. Dieser Indextyp ist verbindlich, wenn die Dokumente, die Sie indexieren, DBCS-Zeichen enthalten. Ein Ngram-Index kann aber auch für SBCS-Dokumente verwendet werden.

Wenn Sie das Schlüsselwort INDEXTYPE nicht angeben, wird der Wert in den Textkonfigurationseinstellungen verwendet.

! Der duale Indextyp kann nicht verwendet werden, um die Vorteile der
! Dokumentstruktur zu nutzen.

Verschachtelte Abschnitte (siehe „Die Vorteile der Struktur eines Dokument nutzen“ auf Seite 68) werden für den Ngram-Indextyp nicht unterstützt.

INDEXOPTION Die bei Erstellung des Indexes zu verwendenden Optionen.

CASE_ENABLED Diese Option ist **nur für Ngram-Indizes** verfügbar. Normalerweise lassen Ngram-Indizes eine Suche unter Beachtung der Groß-/Kleinschreibung nicht zu. Durch die Angabe von CASE_ENABLED stellen Sie sicher, daß Dokumente so indexiert werden, daß eine Suche unter Beachtung der Groß-/Kleinschreibung möglich ist. Weitere Informationen finden Sie in „Ngram-Index“ auf Seite 15.

! **NORMALIZED** Diese Option ist **nur für präzise Indizes** verfügbar. Ein
! normalisierter präziser Index weicht wie folgt von einem präzisen Index ab:

- ! Die Groß-/Kleinschreibung wird nicht beachtet, alle Wörter, mit Ausnahme von Wörtern, die nur aus Großbuchstaben bestehen, werden in Kleinbuchstaben umgesetzt.
- ! Für Wörter, die nur aus Großbuchstaben bestehen, kann das Stoppwortfiltern nicht angewendet werden; die Abkürzung UK wird beispielsweise indexiert.
- ! Suchbegriffe in englischer Sprache können unter Verwendung eines heuristischen Algorithmus erweitert werden, damit sie Grundformen enthalten, so daß eine Suche nach `house` auch nach `houses` sucht.

! **INDEXPROPERTY SECTIONS_ENABLED DOCUMENTMODEL(S) modellname** Merk-
! male eines ausgewählten Indextyps.

! SECTIONS_ENABLED gibt an, daß der ausgewählte Indextyp Informationen zur Dokumentstruktur enthalten kann.

! DOCUMENTMODEL/DOCUMENTMODELS *modellname* gibt das/die
! Modell(e) an, das/die als Standardwert für die zu indexierenden Dokumente zugeordnet werden soll(en). Ein Modellnamen muß angegeben werden, wenn das Indexmerkmal SECTIONS_ENABLED verwendet wird. Wenn eine Liste von Modellen angegeben wird, wird das erste Modell als Standardmodell für den Index verwendet. Das Standardmodell wird während der Indexierung verwendet, wenn das Dokument keinen Verweis auf ein Modell enthält oder wenn während der Suche kein Modell angegeben wird.

! Für den Modellnamen können die Zeichen a - z, A - Z und 0 - 9 verwendet
! werden.

! Der angegebene Modellname muß einer Modelldefinition in der Modelldefinitionsdatei `desmodel.ini` entsprechen.

! Um die Modelle zu ändern, die einem Index zugeordnet sind, verwenden
! Sie

1. ! DISABLE TEXT COLUMN, um den Index zu inaktivieren.
2. ! ENABLE TEXT COLUMN, um die Dokumente erneut zu indexieren,
   ! wobei Sie andere Dokumentmodellnamen angeben.

UPDATEFREQ Aktualisierungsfrequenz Die Indexaktualisierungsfrequenz in bezug auf, wann die Aktualisierung durchgeführt werden soll und wie viele Textdokumente mindestens in der Protokolltabelle stehen müssen. Wenn zum bestimmten Zeitpunkt nicht genügend Textdokumente in der Protokolltabelle stehen, wird der Index nicht aktualisiert.

Die Syntax wird in „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262 beschrieben.

Wenn Sie UPDATEFREQ nicht angeben, wird die Standardfrequenz verwendet, die in den Textkonfigurationseinstellungen angegeben ist.

Hinweis

Wenn Sie viele Tabellen haben, vermeiden Sie möglichst die Verwendung von Standardwerten. Durch das Setzen individueller Aktualisierungsfrequenzeinstellungen für Tabellen können Sie vermeiden, daß alle Tabellen gleichzeitig indexiert werden, wodurch die Systemressourcen unnötig lange belastet werden.

NONE Keine weiteren Indexaktualisierungen werden durchgeführt. Dies ist für eine Textspalte gedacht, an der keine weiteren Änderungen vorgenommen werden.

Diese Einstellungen der Aktualisierungsfrequenz werden ignoriert, wenn sie bereits durch ENABLE TEXT TABLE für die gesamte Tabelle gesetzt wurden.

DIRECTORY verzeichnis Der Verzeichnispfad, in dem der Textindex gespeichert werden soll. Der angegebene Pfad wird verknüpft mit ““txins*nnn*”, wobei *nnn* die Knotennummer ist.

Dies ist ein bestehendes Verzeichnis auf dem System, auf dem der Text Extender-Server läuft.

Wenn Sie das Schlüsselwort DIRECTORY nicht angeben, wird der Wert für DIRECTORY in den Textkonfigurationseinstellungen verwendet.

Diese Einstellung wird ignoriert, wenn sie bereits durch ENABLE TEXT TABLE für die gesamte Tabelle gesetzt wurde.

Wenn Sie partitionierte Datenbanken verwenden, können Sie einen Pfadnamen für den gemeinsamen Index für alle Knoten oder für einen Knotenbereich bzw. unterschiedliche Indexverzeichnispfade für jeden Knoten angeben.

ON NODE knoten [TO knoten] Die Nummer des Knotens oder der Knotenbereich, dem ein Verzeichnispfadname zugeordnet wird.

UPDATEINDEX Ein Schlüsselwort, das festlegt, ob die Textdokumente, die dieser Kennungsspalte zugeordnet sind, sofort nach Abschluß dieses Befehls indexiert werden sollen, ohne auf die nächste regelmäßige Indexierung zu warten, die durch UPDATEFREQ gesetzt ist.

ENABLE TEXT COLUMN (Befehl)

UPDATE Das Indexieren der Textdokumente findet sofort statt, nachdem dieser Befehl abgeschlossen ist.

NOUPDATE Das Indexieren findet zu dem Zeitpunkt statt, der durch die Einstellungen der Aktualisierungsfrequenz entweder in diesem Befehl durch UPDATEFREQ oder durch die Textkonfigurationseinstellung festgelegt ist.

Wenn Sie dieses Schlüsselwort nicht angeben, wird der Wert in den Textkonfigurationseinstellungen verwendet.

COMMITCOUNT anzahl Ein Wert zwischen 500 und 1000000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen DB2 UDB eine sofortige Commit-Anweisung ausgegeben muß. Dadurch können Situationen vermieden werden, in denen nicht genügend Protokollspeicherbereich vorhanden ist, wenn große Tabellen oder Spalten bzw. eine große Anzahl an externen Dateien aktiviert werden.

Hinweise zur Verwendung

Dieser Befehl fügt eine Kennungsspalte zur angegebenen DB2-Tabelle hinzu. Jede Kennungsspalte ist einer Textspalte zugeordnet und wird durch die UDFs des Text Extender verwendet. Wenn diese Tabelle nicht aktiviert wurde, um einen gemeinsamen Index zu erstellen, wird ein Index erstellt, der dieser Textspalte zugeordnet ist.

Außerdem wird eine Protokolltabelle in der Datenbank erstellt. Die Protokolltabelle wird zum Aufzeichnen von Änderungen an der Textspalte verwendet, d.&rbl;h. Einfügungen, Aktualisierungen und Löschungen. Einfügungs-, Aktualisierungs- und Löschauslöser werden für die Textspalte definiert, um die Protokolltabelle automatisch auf dem neuesten Stand zu halten.

Wenn die Textspalte, die Sie aktivieren, zu einer Tabelle gehört, die Teil einer Knotengruppe mit mehreren Knoten ist, muß das Indexverzeichnis, das Sie angeben, auf allen physischen Knoten verfügbar sein. Wenn Sie das Standardverzeichnis verwenden, das in der Textkonfiguration angegeben ist, stellen Sie sicher, daß der Pfad auf allen Knoten der Knotengruppe verfügbar ist. Wenn dies nicht passend ist, können Sie für jeden Knoten einen spezifischen Pfad im Befehl ENABLE TEXT COLUMN angeben. Wenn Sie die Knotenkonfiguration einer Knotengruppe ändern, die eine für den Text Extender aktivierte Tabelle enthält, müssen Sie die Tabelle erneut indexieren.

Hinweis

Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 63.

ENABLE TEXT FILES

Mit Hilfe dieses Befehls kann der Text Extender in Textdateien suchen, die sich nicht in einer DB2 UDB-Datenbank befinden.

Hinweis

Dieser Befehl kann nicht verwendet werden, wenn die Textspalten in der Tabelle einen Index gemeinsam benutzen, wie in „Eine Texttabelle aktivieren“ auf Seite 57 beschrieben.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

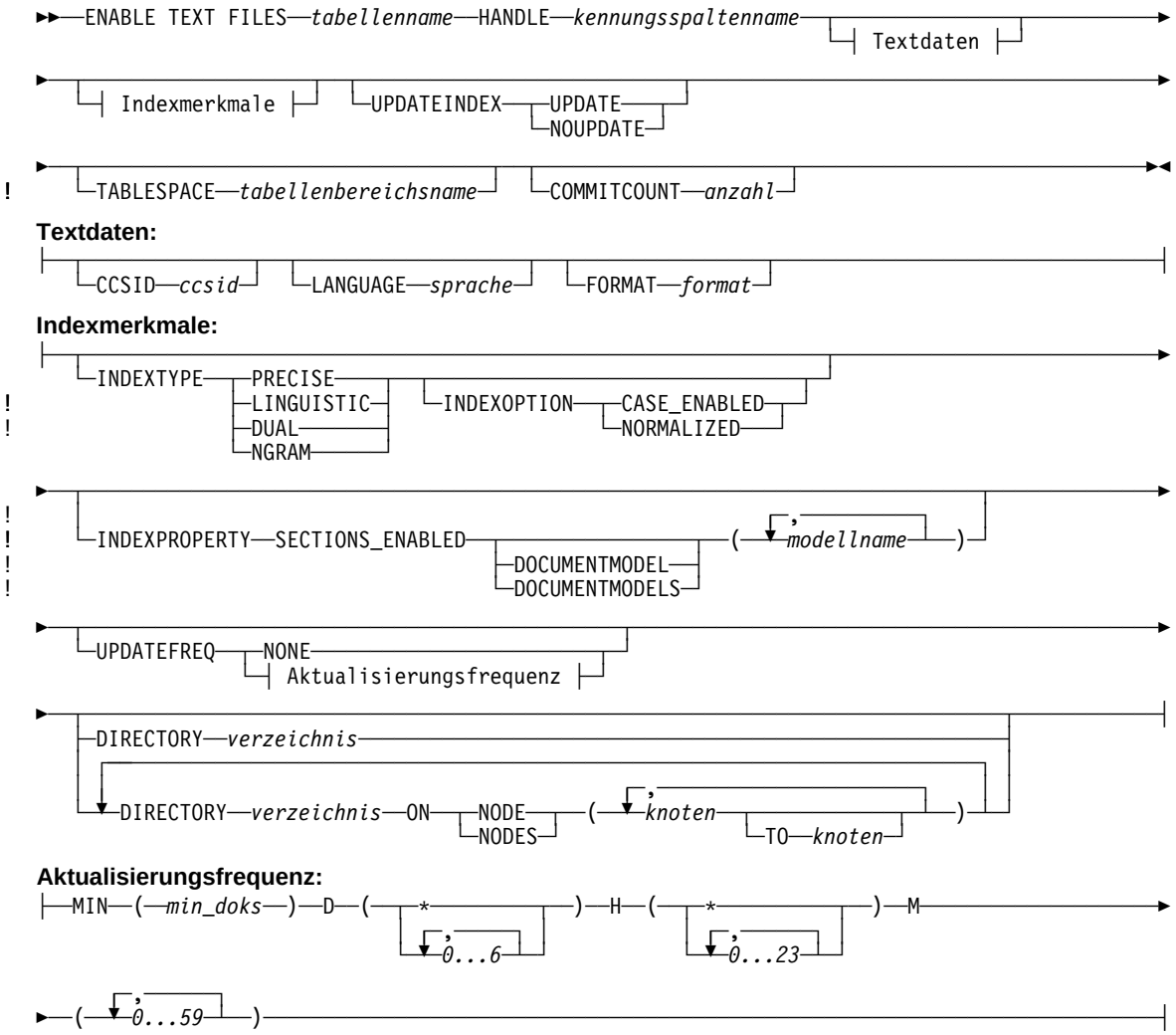
ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

ENABLE TEXT FILES (Befehl)

Befehlssyntax



ENABLE TEXT FILES (Befehl)

Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die den zu indexierenden externen Textdateien zugeordnet werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

Eine Beschreibung der anderen Befehlsparameter finden Sie in „ENABLE TEXT COLUMN“ auf Seite 139.

Hinweise zur Verwendung

Dieser Befehl fügt eine Kennungsspalte zur angegebenen DB2-Tabelle hinzu. Jede Kennungsspalte ist einer Gruppe von externen Textdateien zugeordnet und wird durch die UDFs des Text Extender verwendet. Ein Index wird erstellt, der diesen Dateien zugeordnet ist.

Nachdem Sie die Textdateien aktiviert haben, initialisieren Sie Kennungen in der Kennungsspalte, indem Sie INIT_TEXT_HANDLE verwenden. Füllen Sie anschließend mit Hilfe von UPDATE INDEX den Index.

Sie können einen Kennungsspaltennamen nicht erneut verwenden, wenn dieser Name zuvor im Befehl ENABLE TEXT FILES verwendet wurde, um eine Kennungsspalte einer Textspalte zu identifizieren.

Eine Protokolltabelle wird erstellt, um die Änderungen an den Dateien aufzuzeichnen, aber Sie müssen die Auslöser manuell aktivieren, da DB2 UDB nicht die Kontrolle über die Dateien hat und daher solche Änderungen nicht registriert. Eine Beschreibung dazu finden Sie in „Einen Index für externe Dateien aktualisieren“ auf Seite 78.

Wenn in diesem Schritt der Speicherbereich nicht mehr ausreicht, finden Sie mögliche Lösungen in „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 63.

ENABLE TEXT TABLE (Befehl)

ENABLE TEXT TABLE

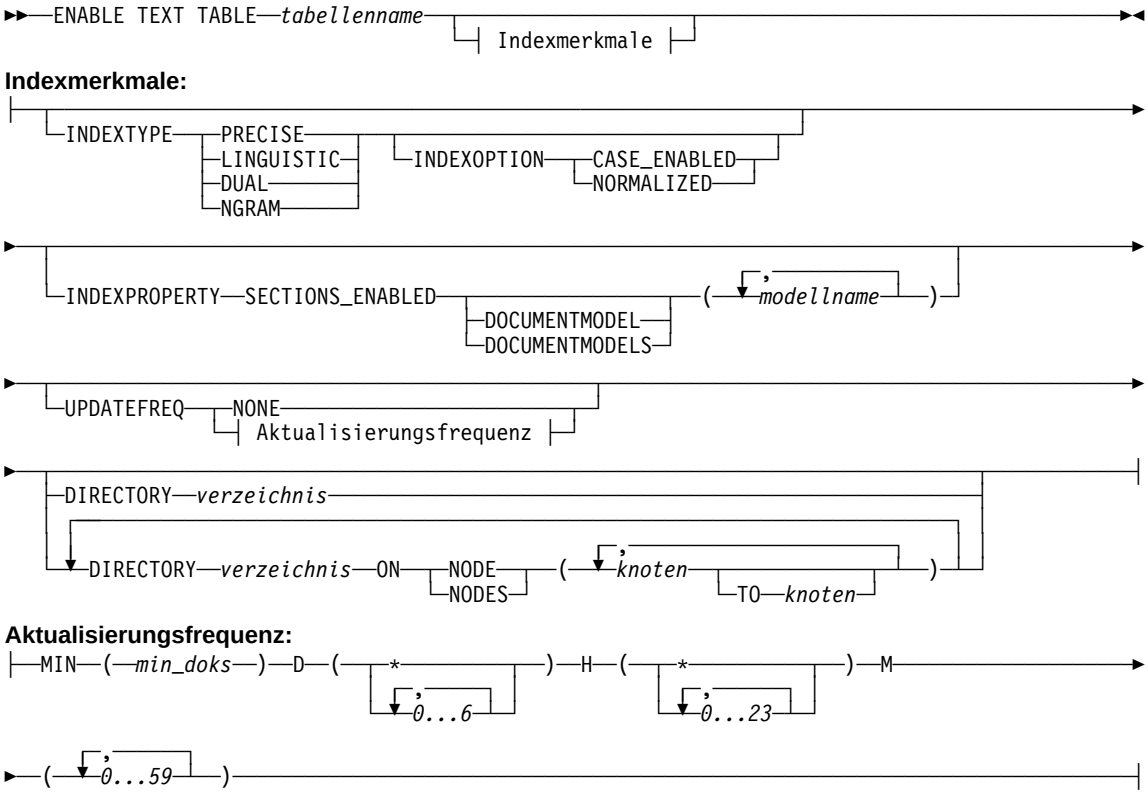
Erstellt einen gemeinsamen Index für die Verwendung von allen Textspalten der Tabelle, die später aktiviert werden. Die Tabelle ist anschließend eine Tabelle mit einem gemeinsam benutzten Index. Eine Tabelle, die nicht auf diese Weise aktiviert wird, sondern in der die Textspalten, die später aktiviert werden, ihre eigenen individuellen Indizes erstellen, ist eine Mehrindextabelle.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung
SELECT-Berechtigung
UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle, die in der verbundenen Datenbank aktiviert werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

INDEXTYPE Der Typ des zu erstellenden Indexes. Weitere Informationen finden Sie in „Indextypen“ auf Seite 13.

PRECISE Die Begriffe werden exakt so indexiert und gesucht, wie sie in den Textdokumenten vorkommen.

LINGUISTIC Die Begriffe werden linguistisch verarbeitet, bevor sie indexiert werden. Die Suchbegriffe werden ebenfalls linguistisch verarbeitet, bevor die Suche beginnt.

DUAL Begriffe werden exakt so indexiert, wie sie in den Textdokumenten vorkommen und sie werden außerdem indexiert, nachdem sie linguistisch verarbeitet wurden. Bei der Suche können Sie für jeden Begriff entscheiden, ob Sie nach dem präzisen Begriff oder nach dem linguistisch verarbeiteten Begriff suchen wollen.

NGRAM Die Begriffe werden indexiert, indem Zeichengruppen syntaktisch analysiert werden, statt ein Wörterverzeichnis zu verwenden. Dieser Wörterverzeichnistyp ist verbindlich, wenn die Dokumente, die Sie indexieren, DBCS-Zeichen enthalten. Ein Ngram-Index kann aber auch für SBCS-Dokumente verwendet werden.

Wenn Sie das Schlüsselwort INDEXTYPE nicht angeben, wird die Textkonfiguration verwendet.

INDEXOPTION Die bei Erstellung des Indexes zu verwendenden Optionen.

CASE_ENABLED Diese Option ist **nur für Ngram-Indizes** verfügbar. Normalerweise lassen Ngram-Indizes eine Suche unter Beachtung der Groß-/Kleinschreibung nicht zu. Durch die Angabe von CASE_ENABLED stellen Sie sicher, daß Dokumente so indexiert werden, daß eine Suche unter Beachtung der Groß-/Kleinschreibung möglich ist. Weitere Informationen finden Sie in „Ngram-Index“ auf Seite 15.

ENABLE TEXT TABLE (Befehl)

! **INDEXPROPERTY SECTIONS_ENABLED DOCUMENTMODEL(S) modellname** Merk-
! male eines ausgewählten Indextyps.

! SECTIONS_ENABLED gibt an, daß der ausgewählte Indextyp Informa-
! tionen zur Dokumentstruktur enthalten kann.

! DOCUMENTMODEL/DOCUMENTMODELS *modellname* gibt das/die
! Modell(e) an, das/die als Standardwert für die zu indexierenden Dokumente
! zugeordnet werden soll(en). Ein Modellnamen muß angegeben werden,
! wenn das Indexmerkmal SECTIONS_ENABLED verwendet wird. Wenn
! eine Liste von Modellen angegeben wird, wird das erste Modell als
! Standardmodell für den Index verwendet. Das Standardmodell wird
! während der Indexierung verwendet, wenn das Dokument keinen Verweis
! auf ein Modell enthält oder wenn während der Suche kein Modell ange-
! geben wird.

! Für den Modellnamen können die Zeichen a - z, A - Z und 0 - 9 verwendet
! werden.

! Der angegebene Modellname muß einer Modelldefinition in der Modelldefi-
! nitionsdatei `desmodel.ini` entsprechen.

! Um die Modelle zu ändern, die einem Index zugeordnet sind, verwenden
! Sie

! 1. DISABLE TEXT TABLE, um den Index zu inaktivieren.

! 2. ENABLE TEXT TABLE, um die Dokumente erneut zu indexieren,
! wobei Sie andere Dokumentmodellnamen angeben.

UPDATEFREQ Aktualisierungsfrequenz Die Indexaktualisierungsfrequenz in bezug auf, *wann* die Aktualisierung durchgeführt werden soll und *wie viele Textdokumente in der Warteschlange* der Protokolltabelle stehen müssen. Wenn zum bestimmten Zeitpunkt nicht genügend Textdokumente in der Protokolltabelle stehen, wird der Index nicht aktualisiert.

Die Syntax wird in „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262 beschrieben.

Wenn Sie UPDATEFREQ nicht angeben, wird die Standardfrequenz verwendet, die in den Textkonfigurationseinstellungen angegeben ist.

NONE Keine weiteren Indexaktualisierungen werden durchgeführt. Dies ist für eine Textspalte gedacht, an der keine weiteren Änderungen vorgenommen werden.

Hinweis

Wenn Sie viele Tabellen haben, vermeiden Sie möglichst die Verwendung von Standardwerten. Durch das Setzen individueller Aktualisierungsfrequenzeinstellungen für Tabellen können Sie vermeiden, daß alle Tabellen gleichzeitig indexiert werden, wodurch die Systemressourcen unnötig lange belastet werden.

DIRECTORY verzeichnis Der Verzeichnispfad, in dem der Textindex gespeichert werden soll. Der angegebene Pfad wird verknüpft mit ""txins*nnn*", wobei *nnn* die Knotennummer ist.

Dies ist ein bestehendes Verzeichnis auf dem System, auf dem der Text Extender-Server läuft.

Wenn Sie das Schlüsselwort DIRECTORY nicht angeben, wird der Wert für DIRECTORY in den Textkonfigurationseinstellungen verwendet.

Wenn Sie partitionierte Datenbanken verwenden, können Sie einen Pfadnamen für den gemeinsamen Index für alle Knoten oder für einen Knotenbereich bzw. unterschiedliche Indexverzeichnispfade für jeden Knoten angeben.

ON NODE knoten [TO knoten] Die Nummer des Knotens oder der Knotenbereich, dem ein Verzeichnispfadname zugeordnet wird.

Hinweise zur Verwendung

Ein neuer Textindex wird erstellt, der allen Textspalten in dieser Tabelle zugeordnet ist. Sie führen diesen Schritt aus, wenn Sie einen gemeinsamen Index für alle Textspalten einer Tabelle haben wollen anstelle von jeweils einem separaten Index für jede Textspalte.

Wenn Sie eine Tabelle aktiviert haben, müssen Sie anschließend ENABLE TEXT COLUMN für jede der Textspalten durchführen, die Sie durchsuchen wollen.

Eine Protokolltabelle wird in der Datenbank erstellt. Die Tabelle wird zum Aufzeichnen von Änderungen, d.&rbl;h. Einfügungen, Aktualisierungen und Löschungen, in den Textspalten verwendet, die später aktiviert werden.

Wenn eine Textspalte aktiviert ist, werden Auslöser erstellt, die Änderungen am Text überwachen und automatisch einen Satz in der Protokolltabelle speichern, welche Dokumente indexiert werden müssen.

Der Text Extender indexiert die Textdokumente, die in der Protokolltabelle aufgelistet sind, regelmäßig, so häufig, wie durch das Schlüsselwort UPDATEFREQ angegeben.

GET INDEX SETTINGS (Befehl)

GET ENVIRONMENT

Dieser Befehl zeigt die Einstellungen der Umgebungsvariablen an.

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

Keine

Hinweise zur Verwendung

Folgende Umgebungsvariablen werden angezeigt:

DB2INSTANCE Der DB2 UDB-Exemplarname.

DB2DBDFT Der Standarddatenbankname.

DB2TX_INSTOWNER (nur UNIX) Der Text Extender-Exemplarname.

DB2TX_INSTOWNERHOMEDIR (nur UNIX) Das Ausgangsverzeichnis des Exemplareigners.

GET INDEX SETTINGS

Dieser Befehl zeigt die folgenden Einstellungen eines Indexes an:

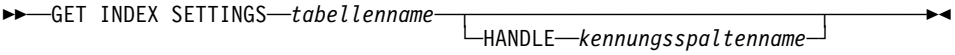
- Indextyp
- Indexoption [wahlfrei]
- Option UPDATE INDEX
- Indexverzeichnis
- Aktualisierungsfrequenz
- Standardmodell

Berechtigung

Keine

GET INDEX SETTINGS (Befehl)

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, deren Indexeinstellungen angezeigt werden sollen. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Indexeinstellungen angezeigt werden sollen.

Hinweise zur Verwendung

Wenn die Tabelle als Mehrindextabelle aktiviert ist, zeigt dieser Befehl die Indexeinstellungen aller aktivierten Textspalten in der Tabelle an. Wenn ein *kennungsspaltenname* geliefert wird, zeigt dieser Befehl die Indexeinstellungen der angegebenen Spalte an.

Wenn die Tabelle eine Tabelle mit einem gemeinsam benutzten Index ist, werden die Einstellungen des gemeinsam benutzten Indexes angezeigt. Wenn ein *kennungsspaltenname* geliefert wird, wird dieser ignoriert.

Wenn die Tabelle oder Spalte mit dem Indexmerkmal SECTIONS_ENABLED aktiviert ist, zeigt der Befehl GET INDEX SETTINGS das Standardmodell für den Index an. Das Standardmodell ist der Modellname, den Sie während der Aktivierung angegeben haben, oder der erste Modellname in einer Liste von Modellnamen. Beispiel:

```
Aktuelle Indexeinstellungen:

      Indextyp                 (INDEXTYPE) = LINGUISTIC
      Standardmodell      (DOCUMENTMODEL)  = mymodel
      Indexaktualisierung    (UPDATEINDEX) = UPDATE
      Aktualisierungsfrequenz (UPDATEFREQ) = NONE
      Knoten 1
      Indexverzeichnis         (DIRECTORY) = user1/db2tx/indexes
```

GET INDEX STATUS (Befehl)

GET INDEX STATUS

Diese Befehl zeigt die folgenden Indexstatusinformationen für eine bestimmte Kennungsspalte oder Tabelle an:

- Ob die Suchfunktion verfügbar ist
- Ob die Indexaktualisierungsfunktion verfügbar ist
- Ob die Reorganisationsfunktion verfügbar ist
- Die Anzahl der geplanten Dokumente
- Die Anzahl der indexierten Dokumente
- Die Anzahl der indexierten Dokumente im Primärindex
- Die Anzahl der indexierten Dokumente im Sekundärindex
- Fehlerereignisse

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Status angezeigt werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Status angezeigt werden soll.

Hinweise zur Verwendung

Für eine Mehrindextabelle müssen Sie den Namen der Kennungsspalte angeben.

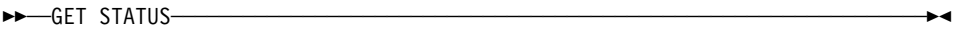
GET STATUS

Dieser Befehl zeigt Informationen zum aktivierten Status von Datenbanken, Tabellen oder Textspalten an.

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

Keine

Hinweise zur Verwendung

Dieser Befehl zeigt an, ob die Datenbank aktiviert ist, zeigt die Namen der aktivierten Texttabellen in der Datenbank, die Namen der aktivierten Textspalten und deren zugeordnete Kennungsspalten und die Spalten von Kennungsspalten für externe Dateien an.

GET TEXT CONFIGURATION

Dieser Befehl zeigt die Standardeinstellungen für die Textkonfiguration für die verbundene Datenbank an.

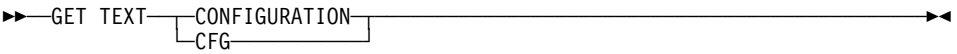
Um diese Standardeinstellungen zu ändern, verwenden Sie „CHANGE TEXT CONFIGURATION“ auf Seite 130.

Berechtigung

Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

SELECT-Berechtigung

Befehlssyntax



Befehlsparameter

Keine

Hinweise zur Verwendung

Ein Beispiel für die Informationen zur Textkonfiguration finden Sie in „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 84.

QUIT (Befehl)

GET TEXT INFO

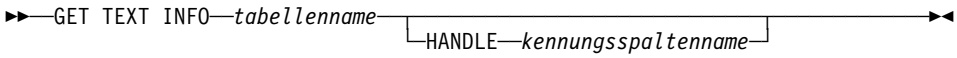
Dieser Befehl zeigt die Textdateneinstellungen für Textspalten an:

- CCSID
- Sprache
- Format

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Textdateneinstellungen angezeigt werden sollen. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Textdateneinstellungen angezeigt werden sollen.

Hinweise zur Verwendung

Wenn der Name einer Kennungsspalte angegeben ist, werden nur die Textdaten für diese Spalte angezeigt.

Wenn kein Name einer Kennungsspalte angegeben ist, werden die Textdaten für jede aktivierte Spalte in dieser Tabelle angezeigt.

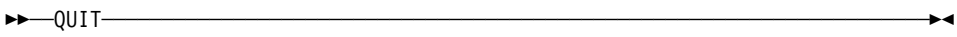
QUIT

Dieser Befehl stoppt den Befehlszeilenprozessor des Text Extender und gibt die Steuerung an das Betriebssystem zurück.

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

Keine

Hinweise zur Verwendung

Die Verbindung zur Datenbank wird beendet.

REORGANIZE INDEX

Wenn eine Textspalte oft aktualisiert wird, wird die Suche im Index ineffizient. Um die Suche wieder effizient zu machen, muß der Index *reorganisiert* werden. Obwohl der Text Extender erkennt, wenn ein Index reorganisiert werden muß und dies automatisch im Hintergrund durchführt, können möglicherweise Situationen auftreten, in denen es erforderlich ist, einen Index manuell unter Verwendung von REORGANIZE INDEX zu reorganisieren. Sie können den Befehl GET INDEX STATUS verwenden, um zu ermitteln, ob ein Index reorganisiert werden muß.

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, deren Index reorganisiert werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Index reorganisiert werden soll.

Hinweise zur Verwendung

Für eine Mehrindextabelle müssen Sie den Namen einer Kennungsspalte angeben.

Obwohl während der Ausführung von REORGANIZE INDEX Suchen durchgeführt werden können, ist dies für Indexaktualisierungen nicht möglich.

UPDATE INDEX (Befehl)

RESET INDEX STATUS

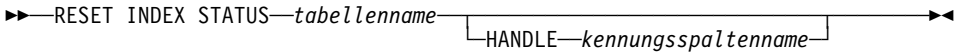
Wenn der Indexstatus einer Tabelle oder Spalte `Suche nicht verfügbar` oder `Aktualisierung nicht verfügbar` anzeigt, ist während der Indexierung ein Fehler aufgetreten, der verhindert, daß Sie mit dem Index arbeiten.

Dieser Befehl setzt den Indexstatus zurück, so daß Sie die Arbeit mit dem Index fortsetzen können. Bevor Sie den Indexstatus zurücksetzen, prüfen Sie mögliche Fehler, die in der Protokolltabelle des Indexes protokolliert wurden (siehe „Fehlerereignisse anzeigen“ auf Seite 86).

Berechtigung

Keine

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalten enthält, deren Status zurückgesetzt werden soll. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Der Name der Kennungsspalte, deren Status zurückgesetzt werden soll.

Hinweise zur Verwendung

Für eine Mehrindextabelle müssen Sie den Namen einer Kennungsspalte angeben.

Für eine Tabelle mit einem gemeinsam benutzten Index wird jede aktivierte Spalte in dieser Tabelle zurückgesetzt.

UPDATE INDEX

Dieser Befehl startet die Indexierung sofort. Er bringt den Index auf den neuesten Stand, um den aktuellen Inhalt der Textspalten widerzuspiegeln, die diesem Index zugeordnet sind.

Damit Aktualisierungen von externen Dateien im Index widergespiegelt werden, erzwingen Sie, daß ein “Änderungseintrag” in die Protokolltabelle gestellt wird, indem Sie eine UPDATE-Anweisung für die entsprechende Kennungsspalte ausgeben. Ein Beispiel finden Sie in „Einen Index für externe Dateien aktualisieren“ auf Seite 78.

UPDATE INDEX (Befehl)

Berechtigung

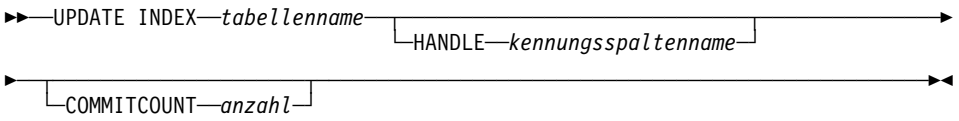
Sie müssen mindestens über eine der folgenden Berechtigungen für die Tabelle verfügen:

ALTER-Berechtigung

SELECT-Berechtigung

UPDATE-Berechtigung

Befehlssyntax



Befehlsparameter

tabellenname Der Name der Texttabelle in der verbundenen Datenbank, die die Textspalte enthält, deren Index aktualisiert werden soll. Dies kann auch der Name einer Tabelle mit einem gemeinsam benutzten Index sein. Der Name muß einen expliziten Schemanamen (ein Qualifikationsmerkmal) enthalten, es sei denn, der Schemaname entspricht Ihrer Benutzer-ID.

HANDLE kennungsspaltenname Wenn dies eine Tabelle mit einem gemeinsam benutzten Index ist, ist der *kennungsspaltenname* nicht erforderlich und wird ignoriert. Der zu aktualisierende Index ist der gesamten Tabelle zugeordnet und nicht einer individuellen Textspalte.

Wenn dies eine Mehrindextabelle ist, ist der *kennungsspaltenname* der Name der Kennungsspalte, deren Index aktualisiert werden soll.

COMMITCOUNT anzahl Ein Wert zwischen 500 und 1000000, der die Anzahl an Einfügungen oder Aktualisierungen angibt, nach denen DB2 UDB eine sofortige Commit-Anweisung ausgegeben muß. Dadurch können Situationen vermieden werden, in denen nicht genügend Protokollspeicherbereich vorhanden ist, wenn große Tabellen oder Spalten bzw. eine große Anzahl an externen Dateien aktiviert werden.

UPDATE INDEX (Befehl)

Kapitel 8. Verwaltungsbefehle für den Server

In diesem Kapitel wird die Syntax von Verwaltungsbefehlen für den Server beschrieben. Die Server-Verwaltung besteht aus Tasks, die Sie zum Starten, Stoppen und Prüfen des Status des Text Extender-Servers und zum Erstellen einer Beispieldatenbank und von Beispieltabellen ausführen können. Kapitel 4, „Verwaltung“ auf Seite 43, beschreibt die Verwendung dieser Befehle.

Befehl	Zweck	Seite
TXICRT	Ein Text Extender-Exemplar erstellen	163
TXIDROP	Ein Text Extender-Exemplar löschen	165
TXILIST (nur UNIX)	Text Extender-Exemplare auflisten (nur UNIX)	165
TXNADD	Text Extender-Server hinzufügen	166
TXNCHECK	Die Text Extender-Datenbankintegrität prüfen	167
TXNDROP	Text Extender-Server löschen	167
TXSAMPLE	Beispieltabellen erstellen und aktivieren	168
TXSTART	Die Text Extender-Services starten	168
TXSTATUS	Den Status des Suchservices anzeigen	169
TXSTOP	Die Text Extender-Services stoppen	169
TXTHESC	Eine Thesaurus-Definitionsdatei kompilieren	169
TXTHESN	Eine Ngram-Thesaurus-Definitionsdatei kompilieren	171
TXTRACE	Trace-Informationen erstellen	172
TXVERIFY	Eine Beispieldatenbank erstellen und aktivieren	178

TXICRT

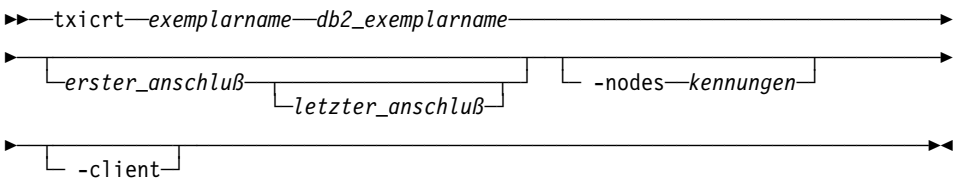
Dieser Befehl erstellt ein Text Extender-Exemplar.

Berechtigung

(Nur UNIX) Root

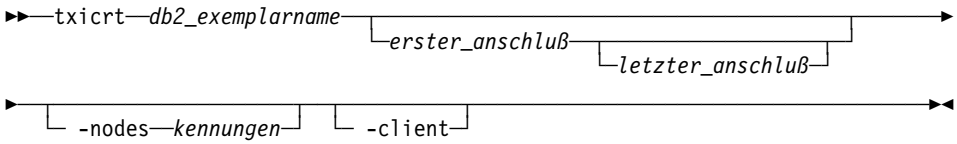
Befehlssyntax

Für UNIX:



TXICRT (Befehl)

Für Windows NT:



Befehlsparameter

exemplarname (Nur UNIX) Die Benutzer-ID des Exemplareigners.

db2_exemplarname Der Name des DB2-Exemplars, das dem zu erstellenden Exemplar zugeordnet werden soll.

erster_anschluß/letzter_anschluß Der TCP/IP-Anschlußbereich, der verwendet werden soll, wenn Sie mit mehr als einem physischen Knoten arbeiten. Dieser Parameter ist nicht erforderlich, wenn Sie nur logischen Knoten verwenden.

Der Anschlußbereich muß mit der Anzahl an Knoten, mit denen Sie arbeiten wollen, übereinstimmen. Wenn Sie mit einem Knoten arbeiten, geben Sie nur einen Parameter `erster_anschluß` an. Um eine Untergruppe von Knoten zu unterstützen, verwenden Sie den Parameter `letzter_anschluß`.

Geben Sie keinen Anschlußbereich an, wenn das Exemplar auf einer Einzelpartitionsdatenbank erstellt wird.

Verwenden Sie nur Anschlußnummern, die nicht bereits in der TCP/IP-Datei `etc/services` aufgelistet sind. Geben Sie keine Anschlußnummern für Installationen an, die nur auf Clients ausgeführt werden.

-nodes *kennungen* Eine Liste von Knotenkennungen. Beispiel: `-nodes 0 1 2`.

Verwenden Sie diesen Parameter, wenn Ihre Tabelle zu bestimmten Knoten in einer Knotengruppe gehört. Wenn dieser Parameter nicht angegeben ist, werden alle Knoten verwendet, d. h. alle Knoten, die in der Datei `db2nodes.cfg` definiert sind.

Geben Sie keine Knotenkennungen für Installationen an, die nur auf Clients ausgeführt werden.

-client

Geben Sie diesen Parameter an, wenn Sie nur eine Installation mit einem Text Extender-Client haben. Wenn Sie diesen Parameter nicht angeben, konfiguriert der Befehl TXICRT ein Client/Server-Exemplar des Text Extender.

Hinweise zur Verwendung

Das Aktivieren von Texttabellen oder Spalten ist nur auf Knoten möglich, auf denen Sie ein Text Extender-Exemplar erstellt haben.

Ein DB2-Exemplar muß bereits für die Benutzer-ID erstellt worden sein.

Bei UNIX-Installationen befindet sich dieser exemplarspezifische Befehl im Verzeichnis `/usr/lpp/db2tx_05_02/instance`.

TXIDROP

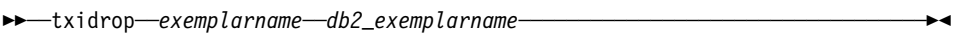
Dieser Befehl löscht ein Text Extender-Exemplar zusammen mit all seinen Indizes.

Berechtigung

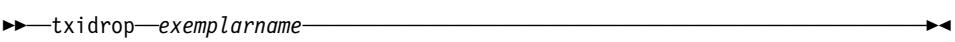
(Nur UNIX) Root

Befehlssyntax

Für UNIX:



Für Windows NT:



Befehlsparameter

exemplarname Der Name des zu löschenden Exemplars.

db2_exemplarname Der Name des DB2-Exemplars, das dem zu löschenden Exemplar zugeordnet ist.

Hinweise zur Verwendung

Bevor Sie ein Exemplar löschen, inaktivieren Sie alle Datenbanken, die für dieses Exemplar aktiviert sind.

Bei UNIX-Installationen befindet sich dieser exemplarspezifische Befehl im Verzeichnis `/usr/lpp/db2tx_05_02/instance`.

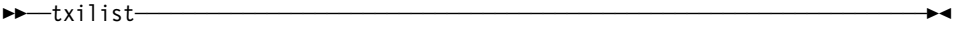
TXILIST (nur UNIX)

Bei UNIX-Systemen listet dieser Befehl die Exemplare für die aktuelle Text Extender-Version auf.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



TXNADD (Befehl)

Befehlsparameter

Keine

Hinweise zur Verwendung

Wenn mehr als eine Version des Text Extender installiert ist, listet dieser Befehl nur die Exemplare für die Text Extender-Version auf, mit der Sie momentan arbeiten, d.&rbl;h. die Version, die durch das aktuelle Verzeichnis bestimmt ist.

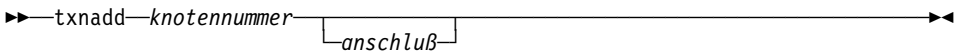
TXNADD

Dieser Befehl erstellt einen zusätzlichen Text Extender-Server auf dem angegebenen Knoten. Durch den Befehl TXSTART wird der Server danach gestartet.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

knotennummer Die Nummer des Knotens, der hinzugefügt werden soll. Geben Sie die Knotennummer exakt so an, wie sie in der Datei `db2nodes.cfg` angegeben ist.

anschluß Die Anschlußnummer, die für den hinzugefügten Knoten verwendet werden soll. Sie müssen eine Anschlußnummer angeben, wenn Sie einen physischen Knoten hinzufügen oder wenn Ihre Knotenkonfiguration physische Knoten enthält.

Hinweise zur Verwendung

Verwenden Sie TXNCHECK, um zu prüfen, welche Indizes Sie erneut erstellen müssen, um synchron mit Ihrer Datenbank zu bleiben.

Hinweis

Wenn Sie einen physischen Knoten zu einer Knotenkonfiguration hinzufügen, die logischen Knoten enthält, müssen Sie zunächst das Exemplar löschen und anschließend das Exemplar erneut erstellen, wobei Sie die Anschlußnummern für alle Knoten angeben.

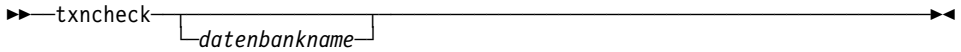
TXNCHECK

Dieser Befehl prüft die Text Extender-Integrität für eine bestimmte Datenbank.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

datenbankname Der Name der zu überprüfenden Datenbank. Wenn kein Datenbankname angegeben ist, wird der Name der DB2-Standarddatenbank verwendet.

Hinweise zur Verwendung

Eine Text Extender-fähige Tabelle muß erneut indexiert werden, wenn die Knotengruppe, in der sich die Tabelle befindet, geändert wird. Verwenden Sie TXNCHECK, um anzuzeigen, welche Text Extender-fähigen Tabellen aufgrund einer neu verteilten Knotengruppe erneut indexiert werden müssen.

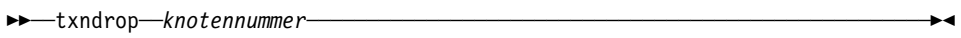
TXNDROP

Dieser Befehl löscht einen Text Extender-Server für einem bestimmten Knoten.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

knotennummer Die Nummer des Knotens, der gelöscht werden soll. Geben Sie die Knotennummer exakt so an, wie sie in der Datei `db2nodes.cfg` angegeben ist.

Hinweise zur Verwendung

Verwenden Sie TXNCHECK, um zu prüfen, welche Indizes Sie erneut erstellen müssen, um synchron mit Ihrer Datenbank zu bleiben.

TXSTART (Befehl)

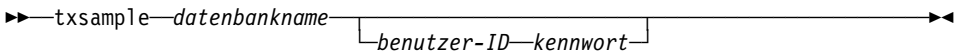
TXSAMPLE

Dieser Befehl erstellt Beispieltabellen, lädt englische Beispieldokumente in die Spalte COMMENT und aktiviert die Textspalte. Siehe „Eine Beispieltabelle erstellen“ auf Seite 54.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

datenbankname Der Name der Datenbank, in der die Beispieltabelle erstellt werden soll.

benutzer-ID Nur erforderlich, wenn Sie an einer Client-Workstation arbeiten.

kennwort Nur erforderlich, wenn Sie an einer Client-Workstation arbeiten.

Hinweise zur Verwendung

Dieser Befehl kann auch auf einer Client-Workstation verwendet werden.

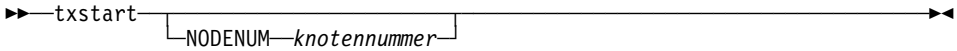
TXSTART

Dieser Befehl startet die Text Extender-Services.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

NODENUM knotennummer Die Knotennummer des zu startenden Servers. Wenn keine Knotennummer angegeben ist, werden alle Server gestartet.

Hinweise zur Verwendung

Führen Sie diesen Befehl in folgenden Situationen aus:

- Während Sie unter einer Benutzer-ID in der SM-Verwaltungsgruppe angemeldet sind.
- Immer, wenn Sie Ihr Server-System stoppen und erneut starten.

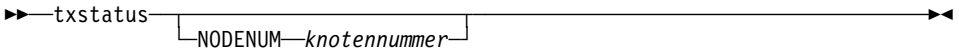
TXSTATUS

Dieser Befehl zeigt an, ob der Text Extender aktiv ist.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

NODENUM knotennummer Die Knotennummer des Servers, dessen Status überprüft werden soll. Wenn keine Knotennummer angegeben ist, wird der Status aller Server überprüft.

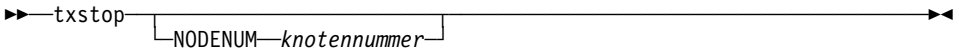
TXSTOP

Dieser Befehl stoppt die Text Extender-Services.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

NODENUM knotennummer Die Knotennummer des zu stoppenden Servers. Wenn keine Knotennummer angegeben ist, werden alle Server gestoppt.

Hinweise zur Verwendung

Mit diesem Befehl wird DB2 nicht gestoppt.

TXTHESC

Dieser Befehl kompiliert eine Thesaurus-Definitionsdatei. Dieser Thesaurus kann nur für Suchen in präzisen oder linguistischen Indizes verwendet werden.

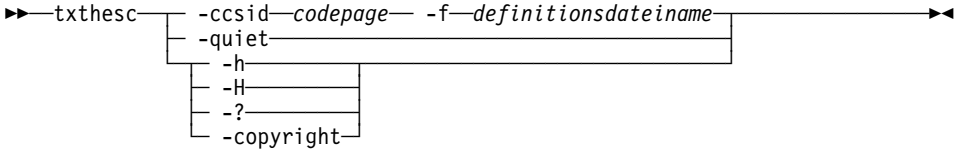
Informationen zum Kompilieren einer Ngram-Thesaurus-Definitionsdatei finden Sie in „TXTHESN“ auf Seite 171

TXTHESC (Befehl)

Berechtigung

(Nur UNIX) Keine

Befehlssyntax



Befehlsparameter

-f *definitionsdateiname* Der Name der SGML-Datei, die die Thesaurus-Definition enthält. Der Dateiname muß entweder den absoluten Pfad oder den relativen Pfad zur Datei enthalten.

Das Thesaurus-Wörterverzeichnis wird in demselben Verzeichnis generiert wie die Definitionsdatei. Es hat den gleichen Namen wie die Definitionsdatei, aber mit den Erweiterungen `th1` bis `th6`.

> **Hinweis**
>
> Da Thesaurus-Dateien überschrieben werden, wenn Sie dieselben Namen haben, sollten Sie jeweils ein separates Verzeichnis für jeden Thesaurus verwenden.

-ccsid *codepage* Die Codepage, in der die Thesaurus-Definitionsdatei geschrieben ist. Momentan wird nur die Codepage 850 unterstützt.

-quiet Die Ausgabeinformationen werden nicht angezeigt.

-copyright Gibt die interne Erstellungsnummer des Produkts zurück. Verwenden Sie diese Nummer beim Berichten von Fehlern.

-h, -H oder -? Zeigt Hilfetext an.

Hinweise zur Verwendung

Verwenden Sie diesen Befehl, um eine Standard-Thesaurus-Definitionsdatei in ein binäres Thesaurus-Definitionsformat zu kompilieren. Die Definitionsdatei muß das SGML-Format haben.

Um eine kompilierte Thesaurus-Datei zu verwenden, verschieben Sie sie in das Verzeichnis für das Wörterverzeichnis des Server-Exemplars und geben Sie anschließend den Standort der Dateien während der Suche an. Das Verzeichnis für das Wörterverzeichnis auf OS/2- und Windows-Systemen ist

laufwerkbuchstabe`:\dmb\db2tx\dict`

Auf AIX-, HP-UX- und SUN-Solaris-Systemen ist das Verzeichnis für das Wörterverzeichnis

DB2TX_INSTOWNER `/db2tx/dicts`

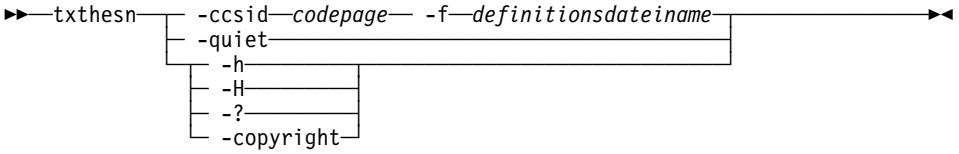
TXTHESN

Dieser Befehl kompiliert eine Ngram-Thesaurus-Definitionsdatei. Dieser Thesaurus kann nur für Suchen in Ngram-Indizes verwendet werden.

Berechtigung

(Nur UNIX) Keine

Befehlssyntax



Befehlsparameter

-f *definitionsdateiname* Der Name der Datei, die die Thesaurus-Definition enthält. Der Dateiname muß entweder den absoluten Pfad oder den relativen Pfad zur Datei enthalten. Der Dateiname ist auf 8 + 3 Zeichen eingeschränkt. Die Erweiterung ist wahlfrei.

Das Thesaurus-Wörterverzeichnis wird in demselben Verzeichnis generiert wie die Definitionsdatei. Sie hat den gleichen Namen wie die Definitionsdatei, aber mit den Erweiterungen `wdf`, `wdv`, `grf`, `grv`, `MEY`, `ROS`, `NEY`, `SOS` und `lk`*n*, wobei *n* eine Ziffer ist.

Hinweis

Da Thesaurus-Dateien überschrieben werden, wenn Sie dieselben Namen haben, sollten Sie jeweils ein separates Verzeichnis für jeden Thesaurus verwenden.

-ccsid *codepage* Die Codepage, in der die Thesaurus-Definitionsdatei geschrieben ist. Eine Liste der unterstützten Codepages finden Sie in „Ngram-Index“ auf Seite 15.

-quiet Die Ausgabeinformationen werden nicht angezeigt.

-copyright Gibt die interne Erstellungsnummer des Produkts zurück. Verwenden Sie diese Nummer beim Berichten von Fehlern.

-h, -H oder -? Zeigt Hilfetext an.

Hinweise zur Verwendung

Verwenden Sie diesen Befehl, um eine Thesaurus-Definitionsdatei in ein binäres Thesaurus-Definitionsformat zu kompilieren. Die Definitionsdatei muß das Format haben, das in „Einen Ngram-Thesaurus erstellen“ auf Seite 41 beschrieben ist.

! Um eine kompilierte Thesaurus-Datei zu verwenden, verschieben Sie sie in das Ver-
! zeichnis für das Wörterverzeichnis des Server-Exemplars und geben Sie anschließend
! den Standort der Dateien während der Suche an. Das Verzeichnis für das Wörterver-
! zeichnis auf OS/2- und Windows-Systemen ist

! *laufwerkbuchstabe*:\dmb\db2tx\dict

! Auf AIX-, HP-UX- und SUN-Solaris-Systemen ist das Verzeichnis für das Wörterver-
! zeichnis

! *DB2TX_INSTOWNER* /db2tx/dicts

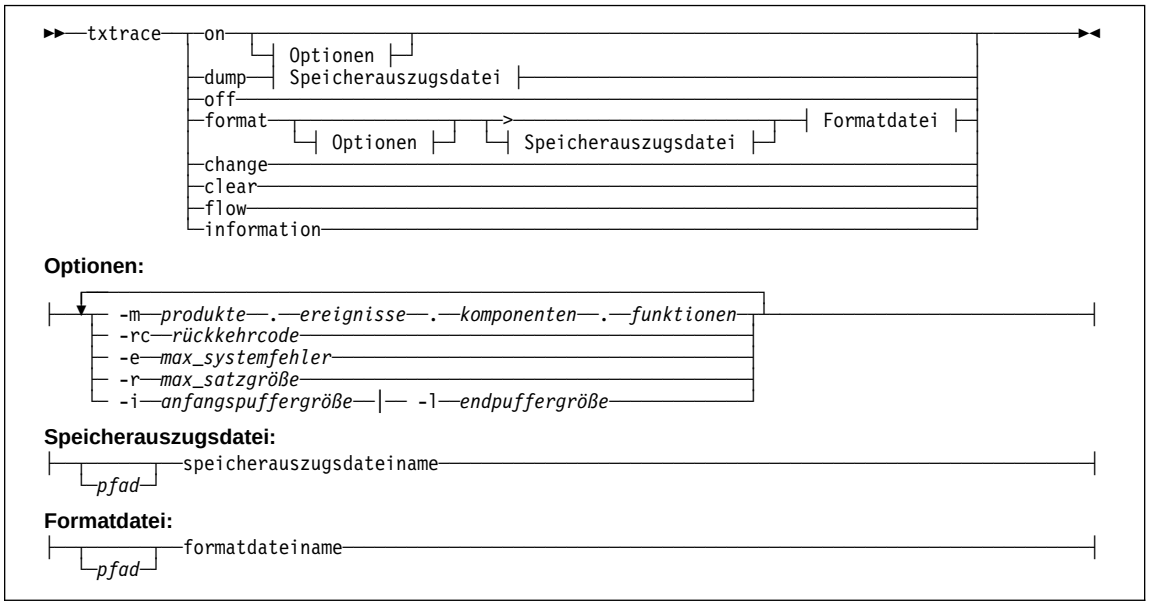
TXTRACE

Dieser Befehl schreibt Verarbeitungsinformationen in einen Trace-Puffer im gemeinsam benutzten Speicher. Diese Informationen können zum Verfolgen von Fehlern verwendet werden. Sie können binär aus dem Trace-Puffer in eine Datei geschrieben werden, um später, wenn die Trace-Funktion ausgeschaltet ist, formatiert zu werden oder sie können formatiert und in eine Datei geschrieben werden, während die Trace-Funktion noch eingeschaltet ist.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



TXTRACE (Befehl)

Befehlsparameter

Anmerkung:

Die Option -u ist auch mit allen TXTRACE-Parametern verfügbar, um Informationen zum Parameter anzuzeigen.

on Zum Starten der Trace-Funktion.

dump | dmp
Zum Schreiben von Trace-Informationen im Binärformat in eine Datei.

off Zum Stoppen der Trace-Funktion.

format | fmt
Zum Formatieren der binären Trace-Informationen. Sie können die Speicherauszugsdatei formatieren, während die Trace-Funktion ausgeschaltet ist, indem Sie den Namen der Speicherauszugsdatei und den Namen der Datei angeben, in der die formatierten Trace-Informationen gespeichert werden sollen. Um die Trace-Informationen direkt aus dem Trace-Puffer zu formatieren, während die Trace-Funktion noch eingeschaltet ist, geben Sie `destrc format > dateiname.tmp` ein.

change | chg
Zum Ändern der Trace-Maske, maxSysErrors oder maxRecordSize.

clear | clr Zum Löschen des Trace-Inhalts.

flow | flw Zum Anzeigen des Steuerungsflusses des Trace.

information | info | inf
Zum Abrufen von Informationen zum Trace.

optionen Zum Filtern der Trace-Informationen, entweder beim Einschalten der Trace-Funktion (dadurch wird die Verwendung des gemeinsam benutzten Speichers verringert) oder beim Formatieren der Trace-Informationen. Es ist normalerweise einfacher, die gesamten Trace-Informationen zu schreiben und sie anschließend bei der Formatierung zu filtern; es sei denn, der Trace ist sehr groß.

-m Zum Hinzufügen einer “Maske”, um anzugeben, welche Ereignisse, Komponenten und Funktionen im Trace enthalten sein sollen. Standardmäßig wird die Trace-Funktion für alles ausgeführt. Die Maske besteht aus vier Teilen, z. B. `2.2-6.1,3.*`, die durch Punkte voneinander getrennt sind. Sie können mit Hilfe des Trennzeichens “-” einen Bereich oder mit Hilfe des Trennzeichens “,” eine Liste angeben. Beispielsweise schließt `2-6` nur Ereignisse ein, deren IDs im Bereich von 2 bis 6 sind. Um nur die Komponenten 2 *und* 6 einzuschließen, geben Sie `2,6` an.

produkte Die Produkt-ID. Die Produkt-ID für den Text Extender ist “2”. Die Produkt-ID für den TextMiner ist “3”.

ereignisse Die folgenden Ereignisarten können in den Trace eingeschlossen werden:

0	system_error
1	system_error
2	system_error
3	non-fatal_error
4	non-fatal_error
5	api_errcode
7	fnc_errcode
8	trap error
10	api_entry
11	api_exit
13	api_retcode
15	api_data
30	fnc_entry
31	fnc_exit
33	fnc_retcode
35	fnc_data

komponenten Die Komponenten, für die der Trace erstellt werden soll.

Die Komponenten-IDs für den Text Extender lauten:

1	COMMAND_LINE_INTERFACE
2	UDF
3	STORED_PROCEDURES
4	ADMINISTRATION
5	INDEX_CONTROL
6	LIBRARY_SERVICES
7	DES_PARSER
8	DES_DEMON
9	DES_API
10	SERVICES

TXTRACE (Befehl)

Die Komponenten-IDs für den TextMiner lauten:

1	automachine
2	bgproc (Hintergrundverarbeitung)
3	cluster
4	common
5	commsrvc (allgemeine Services)
6	communic (Kommunikation)
7	daemon
8	dsclient
9	environ (Umgebung)
10	glue
11	idxcomm (Indexerstellung, allgemeiner Teil)
12	libsrv (Bibliotheksservices)
13	search
14	trace
15	guru
16	indexbld (Indexerstellung, nur tm)
17	indexeng (Indexmaschine, nur tm)
18	smsearch
19	(Suchmaschine, nur tm)
20	tmsearch
21	gtrcm (gtr, allgemeiner Teil)
22	gtrsrch (Suche, nur gtr)
23	gtridx (Indexerstellung, nur gtr)

funktionen Stern (*). Die Gruppe von Funktionen, für die ein Trace durchgeführt werden soll. Verwenden Sie einen Stern (*), um den Trace für alle Funktionen durchzuführen, es sei denn, Sie sind von der IBM Unterstützungsfunktion anders instruiert worden.

-rc rückkehrcode

Behandeln Sie *rückkehrcode* als Systemfehler.

-e max_systemfehler

Eine ganze Zahl. Zum Stoppen des Trace nach dieser Anzahl an Fehlern. Der Standardwert ist 1. Er gibt an, daß wenn der erste Systemfehler auftritt, der nachfolgende Trace von Ereignissen mit niedrigerer Bewertung unterdrückt wird. Dies ist nützlich, wenn Sie nur am ersten Hauptfehler interessiert sind; Sie sollten aber eine höhere Zahl (z.&rbl;B. -e 50) angeben, wenn Sie einen vollständigen Trace nach dem ersten Systemfehler sehen wollen. Der Trace wird im gemeinsam benutzten Speicher gespeichert.

-r max_satzgröße
Eine ganze Zahl. Zum Stoppen des Trace, nachdem diese Anzahl an Sätzen in die Trace-Datei geschrieben wurde. Der Standardwert ist 16 KB.

-i anfangspuffergröße
Eine ganze Zahl. Zum Speichern der Anzahl an Sätzen zu Beginn des Trace. Wenn -i angegeben ist, ist der Standardwert 16 KB. Auf einem UNIX-System ist die empfohlene Puffergröße 2 MB.

Wenn weder -i noch -l angegeben sind, ist -l der Standardwert.

Wenn Sie -i angeben, findet kein Umlauf statt; wenn die Anzahl an Sätzen *max_satzgröße* überschreitet, werden keine weiteren Trace-Einträge geschrieben, auch wenn Sie den Inhalt aller Trace-Einträge löschen. Um neue Trace-Einträge schreiben zu können, erhöhen Sie die Puffergröße, schalten Sie den Trace aus und anschließend wieder ein.

-l endpuffergröße
Eine ganze Zahl. Zum Speichern der Anzahl an Sätzen am Ende des Trace. Der Standardwert ist 16 KB. Auf einem UNIX-System ist die empfohlene Puffergröße 2 MB.

pfad Das Verzeichnis, in dem die entsprechende Datei gespeichert wird.

speicherauszugsdateiname
Der Name der Datei, die die binären Trace-Informationen enthält.

formatierte_datei
Der Name der Datei, die die formatierten Trace-Informationen enthält.

Beispiele

Siehe „Die Trace-Funktion ausführen“ auf Seite 91.

TXVERIFY (Befehl)

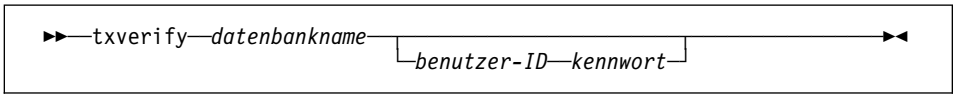
TXVERIFY

Dieser Befehl erstellt eine Beispieldatenbank und aktiviert sie.

Berechtigung

(Nur UNIX) SYSADM, DBADM

Befehlssyntax



Befehlsparameter

datenbankname
Der Name der Datenbank, in der die Beispieltabellen erstellt werden sollen.

benutzer-ID
Nur erforderlich, wenn Sie an einer Client-Workstation arbeiten.

kennwort
Nur erforderlich, wenn Sie an einer Client-Workstation arbeiten.

Hinweise zur Verwendung

Dieser Befehl kann auch auf einer Client-Workstation verwendet werden.

Kapitel 9. UDTs und UDFs

Der Text Extender liefert *benutzerdefinierte Funktionen* (UDFs) zum Suchen in Textdokumenten, die in DB2 UDB gespeichert sind, und zum Arbeiten mit den Ergebnissen einer Suche. Einige der Parameter für UDFs sind Datentypen, die als *benutzerdefinierte eindeutige Typen* (UDTs) bekannt sind und die mit dem Text Extender geliefert werden.

In diesem Kapitel werden die UDTs und die UDFs beschrieben.

Der Text Extender liefert eine DB2 UDB-Eingabedatei für den Befehlszeilenprozessor mit dem Namen `txsample.udf`. Sie enthält Beispiele von Text Extender-UDFs, die mit den Beispieltabellen ausgeführt werden, die in „Die Beispieltabelle DB2TX.SAMPLE“ auf Seite 96 beschrieben werden. Verwenden Sie diese Datei, um Beispiele der Syntax von Verwaltungs- und Such-UDFs anzuzeigen.

Eine Zusammenfassung der Text Extender-UDTs

UDT	Quellendatentyp	Kommentare
DB2TEXTH	VARCHAR(60) FOR BIT DATA	**Textkennung.** Eine Zeichenfolge variabler Länge, die Informationen enthält, die für die Indexierung eines Textdokuments, das in einer Textspalte gespeichert ist, notwendig sind. Zu den Informationen in einer Kennung gehören eine Dokument-ID, der Name des Servers, auf dem der Text indexiert werden soll, der Name des Indexes und Informationen zum Textdokument. Kennungen werden in Spalten gespeichert, die der Text Extender erstellt und der jeweiligen Textspalte zuordnet. Die Funktionen HANDLE und INIT_TEXT_HANDLE geben diesen Datentyp zurück.

Zusammenfassung der UDTs

UDT	Quellendatentyp	Kommentare
DB2TEXTFH	VARCHAR(210) FOR BIT DATA	**Dateikennung.** Eine Zeichenfolge variabler Länge, die Informationen enthält, die für die Indexierung einer externen Textdatei - eine Datei, die außerhalb der Steuerung von DB2 UDB gespeichert ist - notwendig sind. Zu den Informationen in einer Textkennung gehören eine Dokument-ID, der Name des Servers, auf dem der Text indexiert werden soll, der Name des Indexes, Informationen zur Textdatei und zum Standort dieser Datei. Dateikennungen werden in Spalten gespeichert, die der Text Extender erstellt und der jeweiligen Gruppe von externen Dateien zuordnet. Die Funktionen FILE und INIT_TEXT_HANDLE geben diesen Datentyp zurück.
DB2TEXTHLISTP	VARCHAR(16) FOR BIT DATA	**Kennungslistenzeiger.** Ein Zeiger auf eine Liste von Kennungen, die Textdokumenten zugeordnet sind, die in einer Suche gefunden wurden. Die Funktion HANDLE_LIST gibt diesen Datentyp zurück.
DB2TEXTFHLISTP	VARCHAR(16) FOR BIT DATA	**Kennungslistenzeiger.** Ein Zeiger auf eine Liste von Kennungen, die externen Dateien zugeordnet sind, die in einer Suche gefunden wurden. Die Funktion HANDLE_LIST gibt diesen Datentyp zurück.

Eine Zusammenfassung der Text Extender-UDFs

Suchfunktion (UDF)	Zweck	Seite
CCSID	Gibt die CCSID aus der Kennung zurück	182
CONTAINS	Führt eine Suche nach Text in einem bestimmten Dokument durch	182
FILE	Gibt den Pfad und Namen einer Datei in einer bestehenden Kennung zurück oder ändert ihn	183
FORMAT	Gibt die Dokumentformateinstellung in einer Kennung zurück oder ändert sie	184
HANDLE[1]	Gibt eine Kennung aus einer Liste von Kennungen zurück	185
HANDLE_LIST[1]	Sucht eine Liste von Kennungen und gibt sie zurück	186
INIT_TEXT_HANDLE	Gibt eine teilweise initialisierte Kennung zurück, die Informationen, wie z.&rbl;B. Format- und Spracheinstellungen, enthält	187
LANGUAGE	Gibt die Spracheinstellung in einer Kennung zurück oder ändert sie	188
NO_OF_DOCUMENTS[1]	Gibt die Anzahl von Dokumenten zurück, die in einer Kennungsliste aufgelistet sind	189
NO_OF_MATCHES	Sucht die Anzahl der gefundenen Übereinstimmungen und gibt sie zurück	189
RANK	Sucht den Rangordnungswert eines gefundenen Textdokuments und gibt ihn zurück	190
REFINE	Empfängt ein Suchargument und ein Suchargument zur Feinanpassung und gibt ein kombiniertes Suchargument zurück	190
SEARCH_RESULT	Gibt eine Zwischentabelle mit den Suchergebnissen des angegebenen Suchbegriffs zurück	191

In den folgenden Abschnitten werden Syntax und Parameter der UDFs beschrieben. Beispiele zur Verwendung von UDFs werden in Kapitel 5, „Mit Hilfe von Text Extender-UDFs suchen“ auf Seite 95, gegeben.

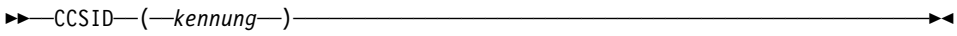
1 Diese Suchfunktionen stammen aus einem früheren Release des Text Extender. Aus Kompatibilitätsgründen werden sie weiterhin unterstützt, ihre Funktionalität wurde durch die PDF SEARCH_RESULT ersetzt.

CONTAINS (UDF)

CCSID

Die Funktion CCSID gibt die ID für codierten Zeichensatz (Datentyp SMALLINT) aus einer Kennung zurück. Dies ist der Parameter CCSID, der für die Indexierung des entsprechenden Textdokuments verwendet wird. Er wird in „CCSIDs“ auf Seite 259 beschrieben. Er wird durch den Befehl ENABLE TEXT COLUMN für jede Textspalte gesetzt.

Funktionssyntax



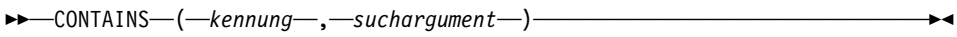
Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, aus der die CCSID-Einstellung zurückgegeben werden soll.

CONTAINS

Die Funktion CONTAINS sucht nach Text in einem bestimmten Textdokument. Sie gibt den INTEGER-Wert 1 zurück, wenn das Dokument den Text enthält. Andernfalls gibt sie den Wert 0 zurück.

Funktionssyntax



Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, die die Kennungen von zu durchsuchenden Textdokumenten enthält.

suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

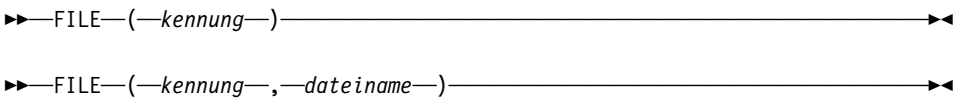
FILE

Die Funktion FILE führt eine der folgenden Aktionen aus:

- Sie gibt den Pfad und Dateinamen in einer Kennung zurück.
- Sie ändert den Pfad und Dateinamen in einer Kennung und gibt den Pfad und Dateinamen zurück.

Die zurückgegebene Kennung ist ein Wert vom Typ DB2TEXTFH.

Funktionssyntax



Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, aus der der Dateiname zurückgegeben werden soll.

dateiname Eine Zeichenfolge vom Typ VARCHAR(150), die den neuen absoluten Pfad und Dateinamen der externen Datei angibt, die der Kennung zugeordnet werden soll. Der Pfad könnte beispielsweise ein LAN-Laufwerk oder ein als NFS angehängtes Laufwerk sein. Die Dateizugriffsberechtigungen müssen den Zugriff auf die Datei von dem DB2 UDB-Exemplareigner zulassen.

FORMAT (UDF)

FORMAT

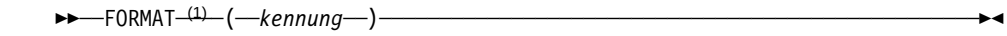
Die Funktion FORMAT führt eine der folgenden Aktionen aus:

- Sie gibt das in einer Kennung angegebene Dokumentformat zurück.
- Sie ändert die Formatangabe in der Kennung eines Dokuments und gibt die geänderte Kennung zurück.

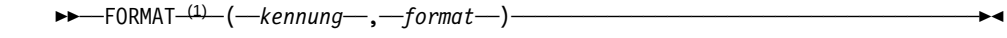
Das zurückgegebene Dokumentformat ist eine Zeichenfolge vom Typ VARCHAR(30). Die zurückgegebene Kennung ist vom Typ DB2TEXTH oder DB2TEXTFH.

Dies ist der Formatparameter, der für die Indexierung des entsprechenden Textdokuments verwendet wird. Die unterstützten Dokumentformate werden in „Formate“ auf Seite 256 aufgelistet.

Funktionssyntax



Anmerkung:
1 Gibt einen Formatwert vom Typ VARCHAR(30) zurück.



Anmerkung:
1 Gibt eine Kennung, Typ DB2TEXTH oder DB2TEXTFH zurück.

Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, aus der die Formateinstellung zurückgegeben werden soll oder in der sie gesetzt werden soll.

format Die neue Einstellung für das Dokumentformat vom Datentyp VARCHAR(30).

Wenn *format* angegeben ist, wird dieses Dokumentformat in der Kennung gesetzt; in diesem Fall wird die Kennung anstelle der Formateinstellung zurückgegeben.

HANDLE

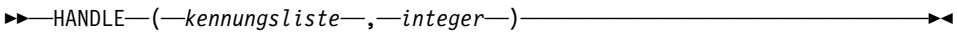
Die Funktion HANDLE gibt eine Kennung vom Typ DB2TEXTH oder DB2TEXTFH zurück, die nach Folgenummer aus einer Liste von Kennungen ausgewählt wird.

Verwenden Sie die Funktion HANDLE nur in einer INSERT-Anweisung.

! **Hinweis**

! Diese Funktion stammt aus einem früheren Release des Text Extender. Aus Kom-
! patibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde durch die
! PDF SEARCH_RESULT ersetzt.

Funktionssyntax



Funktionsparameter

kennungsliste Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTHLISTP oder DB2TEXTFHLISTP ist. Er wird durch die Funktion HANDLE_LIST zurückgegeben.

Dies ist ein Zeiger auf eine Liste von Kennungen von Dokumenten, die in einer Suche gefunden wurden.

integer Ein INTEGER-Wert, der angibt, welche Kennung in der Liste zurückgegeben werden soll.

HANDLE_LIST

Die Funktion HANDLE_LIST sucht unter Verwendung eines Sucharguments nach Textdokumenten und gibt einen Wert vom Typ DB2TEXTHLISTP oder DB2TEXTFHLISTP zurück, der auf eine Liste von Kennungen für die gefundenen Dokumente zeigt.

Anmerkung: Auf den Zeiger auf diese Liste kann nur innerhalb des Bereichs von SQL-Anweisungen zugegriffen werden, die diese Funktion verwenden.

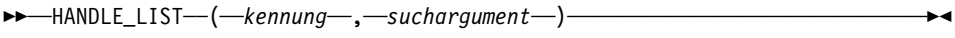
Wenn keine Textdokumente mit den Suchbegriffen gefunden werden, ist die Liste von Kennungen leer. Verwenden Sie die Funktion NO_OF_DOCUMENTS, um festzustellen, ob die Liste leer ist.

Wenn Sie nach Text suchen, der in einer Spalte einer Basistabelle gespeichert ist, verwenden Sie diese Funktion nur einmal pro Anweisung. Wenn Sie nach Text suchen, der in einer Spalte einer Sicht gespeichert ist, und die Sicht aus mehr als einer Basistabelle erstellt wurde (d.&rbl;h. eine Gesamtverknüpfungssicht), verwenden Sie eine Kennung für jeweils eine Kennungsspalte als Eingabeargument *kennung* für die Funktion HANDLE_LIST.

Hinweis

Diese Funktion stammt aus einem früheren Release des Text Extender. Aus Kompatibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde durch die PDF SEARCH_RESULT ersetzt.

Funktionssyntax



Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise die Kennungsspalte, die die Kennungen von zu durchsuchenden Textdokumenten enthält. Es kann auch eine Prototypkennung sein, die aus der Katalogsicht TEXTCOLUMNS extrahiert wurde.

suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

INIT_TEXT_HANDLE

Die Funktion INIT_TEXT_HANDLE gibt eine teilweise initialisierte Kennung zurück, die vordefinierte Werte für Format oder Sprache des Textes enthält. Sie kann in eine Kennungsspalte eingefügt werden. Dies ist sinnvoll, wenn Sie eine Zeile hinzufügen, die Text enthält, dessen Format und Sprache nicht mit den Angaben in den Textkonfigurationseinstellungen übereinstimmt.

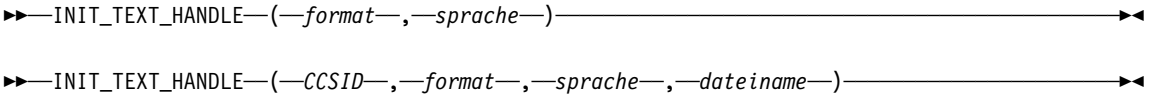
Die zurückgegebene Kennung ist ein Wert vom Typ DB2TEXTH.

Wenn Sie beabsichtigen, Text zu durchsuchen, der in *externen Dateien* gespeichert ist statt in einer DB2 UDB-Tabelle, können Sie die Funktion INIT_TEXT_HANDLE verwenden, um eine vollständig initialisierte Kennung zurückzugeben, die vordefinierte Werte für CCSID, Format und Sprache des Textes und den Standort der Datei enthält.

Die zurückgegebene Kennung ist ein Wert vom Typ DB2TEXTFH.

Verwenden Sie die Funktion INIT_TEXT_HANDLE, um Kennungswerte einzufügen oder zu aktualisieren.

Funktionssyntax



Funktionsparameter

format Eine Zeichenfolge vom Typ VARCHAR(30), die die neue Einstellung für das Dokumentformat angibt. Die unterstützten Formate werden in „Formate“ auf Seite 256 aufgelistet.

sprache Eine Zeichenfolge vom Typ VARCHAR(30), die die neue Einstellung für die Dokumentsprache angibt. Die unterstützten Sprachen werden in „Sprachen“ auf Seite 258 aufgelistet.

dateiname Eine Zeichenfolge vom Typ VARCHAR(150), die den absoluten Pfad und Dateinamen der externen Datei angibt, die der Kennung zugeordnet werden soll. Um unter UNIX Zugriff auf die Datei zu haben, muß der DB2 UDB-Exemplareigner in den Dateizugriffsberechtigungen enthalten sein. Für OS/2- und Windows-Benutzer müssen die Dateizugriffsberechtigungen die Anmeldebenutzer-IDs enthalten.

LANGUAGE (UDF)

LANGUAGE

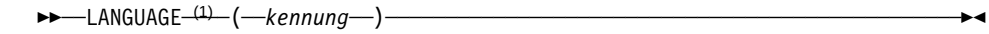
Die Funktion LANGUAGE führt eine der folgenden Aktionen aus:

- Sie gibt die in einer Kennung angegebene Dokumentsprache zurück.
- Sie ändert die Sprachangabe in der Kennung eines Dokuments und gibt die geänderte Kennung zurück.

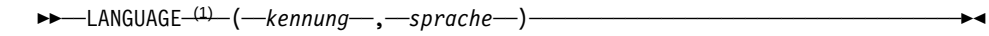
Die zurückgegebene Dokumentsprache ist eine Zeichenfolge vom Typ VARCHAR(30). Die zurückgegebene Kennung ist vom Typ DB2TEXTH oder DB2TEXTFH.

Dies ist der Sprachparameter, der für die Indexierung des entsprechenden Textdokuments verwendet wird. Die unterstützten Sprachen werden in „Sprachen“ auf Seite 258 aufgelistet.

Funktionssyntax



Anmerkung:
1 Gibt einen Sprachwert vom Typ VARCHAR(30) zurück.



Anmerkung:
1 Gibt eine Kennung vom Typ DB2TEXTH oder DB2TEXTFH zurück.

Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, aus der die Spracheinstellung zurückgegeben werden soll oder in der sie gesetzt werden soll.

sprache Die neue Einstellung für die Dokumentsprache vom Datentyp VARCHAR(30).

Wenn *sprache* angegeben ist, wird diese Dokumentsprache in der Kennung gesetzt; die Kennung wird anstelle der Spracheinstellung zurückgegeben.

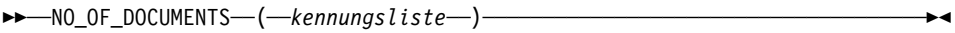
NO_OF_DOCUMENTS

Die Funktion NO_OF_DOCUMENTS gibt einen INTEGER-Wert zurück, der die Anzahl von Elementen in einer Liste von Textdokumenten, die in einer Suche gefunden wurden, angibt. Der zurückgegebene Wert ist die Anzahl von Einträgen, die in einer Liste von Kennungen gefunden wurden.

Hinweis

Diese Funktion stammt aus einem früheren Release des Text Extender. Aus Kompatibilitätsgründen wird sie weiterhin unterstützt, ihre Funktionalität wurde durch die PDF SEARCH_RESULT ersetzt.

Funktionssyntax



Funktionsparameter

kennungsliste Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTHLISTP oder DB2TEXTFHLISTP ist. Er wird durch die Funktion HANDLE_LIST zurückgegeben.

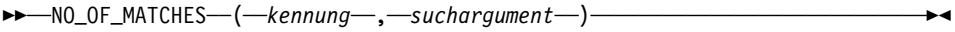
Dies ist ein Zeiger auf eine Liste von Kennungen von Dokumenten, die in einer Suche gefunden wurden.

Die UDFs HANDLE_LIST und NO_OF_DOCUMENTS müssen in derselben SQL-Anweisung angegeben sein, da die Liste nur innerhalb des Bereichs der Anweisung existiert.

NO_OF_MATCHES

NO_OF_MATCHES kann in Textdokumenten suchen und einen INTEGER-Wert zurückgeben, der angibt, wieviele Übereinstimmungen pro Dokument als Ergebnis ermittelt wurden.

Funktionssyntax



REFINE (UDF)

Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, die die Kennungen von zu durchsuchenden Textdokumenten enthält.

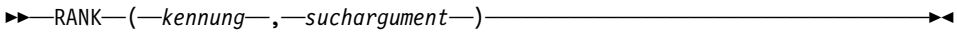
suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

RANK

RANK kann in Textdokumenten suchen und gibt einen Wert für jedes gefundene Dokument zurück, der angibt, wie gut das gefundene Dokument durch das Suchargument beschrieben ist.

RANK gibt einen DOUBLE-Wert zwischen 0 und 1 zurück. Der Rangordnungswert ist absolut und gibt an, wie gut das gefundene Dokument die Suchkriterien in Relation zu anderen gefundenen Dokumenten erfüllt. Der Wert gibt die Anzahl der im Dokument gefundenen Übereinstimmungen in Relation zur Größe des Dokuments an.

Funktionssyntax



Funktionsparameter

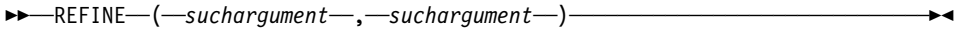
kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, die die Kennungen von zu durchsuchenden Textdokumenten enthält.

suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

REFINE

Die Funktion REFINE empfängt zwei Suchargumente und gibt ein kombiniertes Suchargument vom Typ LONG VARCHAR zurück, das aus den zwei ursprünglichen Suchargumenten besteht, die durch den Booleschen Operator AND verknüpft sind.

Funktionssyntax



Funktionsparameter

suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

SEARCH_RESULT

Die Funktion SEARCH_RESULT gibt das Ergebnis einer Suche in einer Zwischentabelle zurück. Diese UDF kann in einer FROM-Klausel einer SQL-Anweisung verwendet werden.

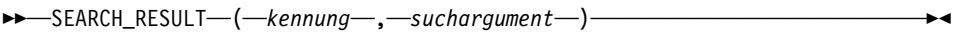
Die zurückgegebenen Tabelle hat die folgende Struktur:

Spaltenname	Datentyp
HANDLE	DB2TX.DB2TEXTH, DB2TX.DB2TEXTFH
NUMBER_OF_MATCHES	INTEGER
RANK	DOUBLE

Werte werden nur für die ausgewählten Spalten der Zwischentabelle generiert. SELECT COUNT(*) generiert nur die Spalte HANDLE. Da die Berechnung der Rangordnungswerte sehr viel Systemressourcen verbraucht, sollten Sie den Rangordnungswert für die Zwischentabelle nicht auswählen, wenn der Rangordnungswert nicht erforderlich ist.

Diese UDF ist schneller als CONTAINS oder RANK, wenn große Tabellen verarbeitet werden.

Funktionssyntax



Funktionsparameter

kennung Ein Ausdruck, dessen Ergebnis ein Wert vom Typ DB2TEXTH oder DB2TEXTFH ist. Normalerweise der Name einer Kennungsspalte, die die Kennungen von zu durchsuchenden Textdokumenten enthält.

suchargument Eine Zeichenfolge vom Typ LONG VARCHAR, die die Begriffe enthält, nach denen gesucht werden soll. Siehe Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193.

Beispiele

Ein Beispiel finden Sie in „Die Suchleistung verbessern“ auf Seite 115. Außerdem können Sie sich die Beispiel-UDFs ansehen, die in der Datei geliefert werden, die in „Die Beispiel-UDFs“ auf Seite 95 beschrieben ist.

SEARCH_RESULT (UDF)

Kapitel 10. Syntax von Suchargumenten

Ein Suchargument ist eine Bedingung, die Sie angeben, wenn Sie nach Begriffen in Textdokumenten suchen. Es besteht aus einem oder mehreren Suchbegriffen und Suchparametern.

Beispiele von Suchargumenten werden in „Suchargumente angeben“ auf Seite 102 und in der Datei `txsample.udf` aufgeführt. Darin sind Beispiele von Text Extender-UDFs enthalten, die mit der Beispieltabelle ausgeführt werden, die in „Die Beispieltabelle DB2TX.SAMPLE“ auf Seite 96 beschrieben ist.

Zu den UDFs, die Suchargumente verwenden, gehören:

- CONTAINS. Diese Funktion verwendet ein Suchargument, um in einem bestimmten Textdokument nach Text zu suchen. Sie gibt den INTEGER-Wert 1 zurück, wenn das Dokument den Text enthält. Andernfalls gibt sie den Wert 0 zurück.
- NO_OF_MATCHES. Diese Funktion verwendet ein Suchargument, um in Textdokumenten zu suchen. Sie gibt einen INTEGER-Wert zurück, der angibt, wie viele Übereinstimmungen pro Dokument als Ergebnis ermittelt wurden.
- RANK. Diese Funktion verwendet ein Suchargument, um in Textdokumenten zu suchen. Sie gibt einen Wert für jedes gefundene Dokument zurück, der angibt, wie gut das gefundene Dokument durch das Suchargument beschrieben ist.
- REFINE. Diese Funktion empfängt zwei Suchargumente und gibt ein kombiniertes Suchargument vom Typ LONG VARCHAR zurück, das aus den zwei ursprünglichen Suchargumenten besteht, die durch den Booleschen Operator AND verknüpft sind.
- HANDLE_LIST. Diese Funktion verwendet ein Suchargument, um in Textdokumenten zu suchen. Sie gibt einen Wert vom Typ DB2TEXTHLISTP oder DB2TEXTFHLISTP zurück, der auf eine Liste von Kennungen für die gefundenen Dokumente zeigt.
- SEARCH_RESULT. Diese Funktion gibt eine Tabelle zurück, die die angeforderten Informationen, d.&rbl;h. die Rangordnung, die Anzahl an Übereinstimmungen und die Kennung, enthält.

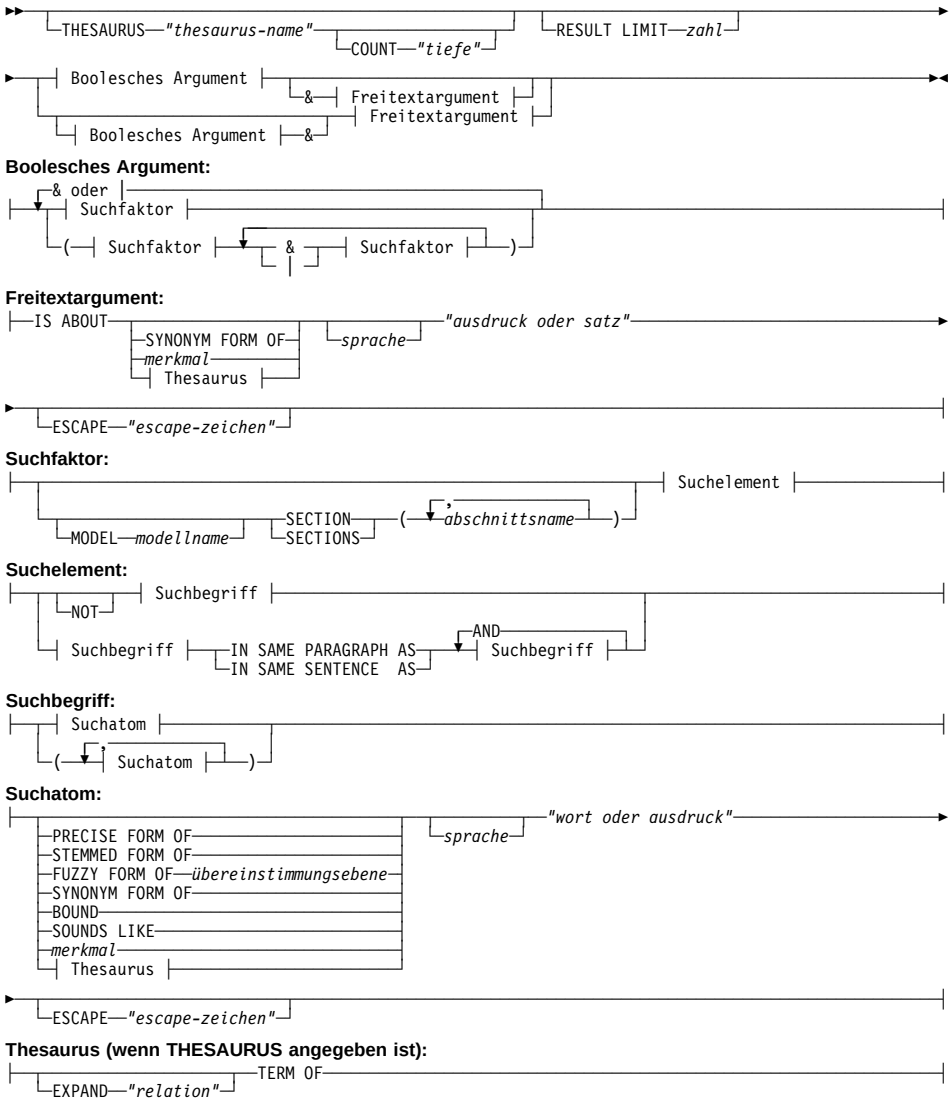
Zu den API-Funktionen, die Suchargumente verwenden, gehören:

- DesGetBrowseInfo. Diese Funktion verwendet ein Suchargument für die Suche in Text, der durch eine Kennung identifiziert ist. Sie gibt einen Zeiger auf Anzeigeinformationen zurück, die von DesStartBrowseSession zur Hervorhebung der Begriffe benötigt werden.
- DesGetSearchResultTable. Diese Funktion verwendet ein Suchargument für die Suche in Textdokumenten, die durch eine Textspalte identifiziert sind. Die Kennungsdaten der gefundenen Textelemente werden in eine Ergebnistabelle geschrieben. Anzeigeinformationen zur Rangordnung und der Anzahl der Übereinstimmungen können ebenfalls in die Ergebnistabelle geschrieben werden.

Suchargument

Suchargument

Suchargumentsyntax



Beispiele

Beispiele werden in „Suchargumente angeben“ auf Seite 102 geliefert.

Suchparameter

IS ABOUT

Eine Option, mit der Sie ein Freitextsuchargument angeben können, d.&rbl;h. ein Ausdruck oder ein Satz in natürlicher Sprache, der das zu suchende Konzept beschreibt. Siehe „Freitext- und Hybridsuche“ auf Seite 111.

! **MODEL** *modellname* Ein Schlüsselwort, mit dem der Name des Dokumentmodells
! angegeben wird, das im Suchbegriff verwendet werden soll. Das
! Dokumentmodell beschreibt die Struktur von Dokumenten, die
! identifizierbare Abschnitte enthalten, so daß der Inhalt dieser Abschnitte
! individuell durchsucht werden kann.

! Der Modellname muß in einer Dokumentmodelldatei angegeben werden,
! die in „Die Vorteile der Struktur eines Dokument nutzen“ auf Seite 68
! beschrieben ist. Der Modellname kann mit Platzhalterzeichen angegeben
! werden.

! Wenn Sie kein Modell angeben, wird das Standardmodell verwendet, das
! während der Indexerstellung angegeben wurde.

! **SECTION(S)** *abschnittsname* Ein Schlüsselwort, mit dem ein oder mehrere Abschnitte
! angegeben werden, auf die die Suche beschränkt werden soll. Der
! Abschnittsname muß in einem Modell in einer Dokumentmodelldatei ange-
! geben werden, die in „Die Vorteile der Struktur eines Dokument nutzen“ auf Seite 68 beschrieben ist. Ein Abschnittsname kann mit Platzhalterzeichen
! angegeben werden. Abschnitte können innerhalb von anderen Abschnitten
! verschachtelt sein. Beispiel:

! `play/Act/Title=play/act/title`

! **Einschränkungen:** Die Suche in verschachtelten Abschnitten ist nur in
! XML-Dokumenten möglich. Bei Ngram-Indizes kann nur der Name des
! Abschnitts gesucht werden und dieser darf kein verschachtelter Abschnitt
! sein.

Suchargument

! **THESAURUS** *thesaurus-name* Ein Schlüsselwort, mit dem der Name des Thesaurus
! angegeben wird, der zum Erweitern des Suchbegriffs verwendet werden
! soll. Der Thesaurus-Name ist der Name (ohne Erweiterung) einer
! Thesaurus-Datei, die unter Verwendung des Thesaurus-Compilers
! TXTHESC oder TXTHESN kompiliert wurde. Es gibt die Standard-Thesauri
! `desthes` und `desnthes`, die im Beispielverzeichnis gespeichert sind, wobei
! `desnthes` ein Ngram-Thesaurus ist. Sie können auch den Pfadnamen der
! Datei angeben. Der Standardpfadname ist der Wörterverzeichnispfad.

COUNT *tiefe* Ein Schlüsselwort, mit dem die Anzahl der Ebenen (die Tiefe) von Begriffen im Thesaurus angegeben wird, die verwendet werden sollen, um den Suchbegriff für eine bestimmte Relation zu erweitern. Wenn Sie dieses Schlüsselwort nicht angeben, wird von einer COUNT-Angabe von 1 ausgegangen.

RESULT LIMIT *zahl* Ein Schlüsselwort, mit dem die maximale Anzahl von Einträgen angegeben wird, die in der Ergebnisliste zurückgegeben werden sollen. *zahl* ist ein Wert von 1 bis 32767. Wenn eine Freitextsuche verwendet wird, wird die Suchergebnisliste nur bezüglich der vollständigen Suchergebnisliste in eine Rangordnung gebracht. Andernfalls werden beim begrenzten Suchergebnis die Einträge der Liste in eine Rangordnung gebracht.

EXPAND *relation* Ein Schlüsselwort, mit dem die Relation, z.&rbl;B. INSTANCE, zwischen dem in TERM OF angegebenen Suchbegriff und den Thesaurus-Begriffen angegeben wird, die zum Erweitern des Suchbegriffs verwendet werden sollen. Der Relationsname muß mit einer Relation übereinstimmen, die im Thesaurus verwendet wird. Siehe „Thesaurus-Konzepte“ auf Seite 32.

TERM OF *"wort oder ausdruck"* Der Suchbegriff oder Mehrwortsuchbegriff, zu dem andere Suchbegriffe aus dem Thesaurus hinzugefügt werden sollen.

Suchargument

Suchfaktor Ein Operand, der mit anderen Operanden kombiniert werden kann, um ein Suchargument zu bilden. Die Auswertung erfolgt von links nach rechts. Der logische Operator AND (&) bindet stärker als der logische Operator OR (|). Beispiel:

```
"passenger" & "vehicle" | "transport" & "public"
```

wird ausgewertet als

```
("passenger" & "vehicle") | ("transport" & "public")
```

Zum Suchen nach

```
"passenger" & ("vehicle" | "transport") & "public"
```

müssen Sie runde Klammern wie gezeigt einfügen.

NOT suchbegriff Ein Operator, mit dem Sie Textdokumente, die einen bestimmten Begriff enthalten, ausschließen können.

Wenn NOT in einem Suchfaktor verwendet wird, können Sie das Schlüsselwort SYNONYM FORM OF nicht verwenden.

suchbegriff IN SAME PARAGRAPH AS suchbegriff Ein Schlüsselwort, mit dem Sie nach einer Kombination von Begriffen, die im selben Absatz auftreten, suchen können.

Die folgenden Suchargumente finden Textdokumente, die den Begriff “traffic” enthalten, nur, wenn der Begriff “air” im selben Absatz auftritt.

```
"traffic" IN SAME PARAGRAPH AS "air"
```

Sie können das Schlüsselwort IN SAME PARAGRAPH AS nicht verwenden, wenn NOT in einem Suchfaktor verwendet wird.

suchbegriff IN SAME SENTENCE AS suchbegriff Ein Schlüsselwort, mit dem Sie nach einer Kombination von Begriffen, die im selben Satz auftreten, suchen können. Ähnlich wie IN SAME PARAGRAPH AS.

AND suchbegriff Ein Schlüsselwort, mit dem Sie mehrere Suchbegriffe kombinieren können, nach denen im selben Satz oder Absatz gesucht werden soll.

Die folgenden Suchargumente suchen nach “forest”, “rain”, “erosion” and “land” im selben Satz.

```
"forest" IN SAME SENTENCE AS "rain" AND "erosion" AND "land"
```

Suchargument

suchatom Wenn Sie eine Serie von Suchatomen durch Kommas verbinden, ist eine Suche erfolgreich, wenn ein Begriff in einem der Suchatome gefunden wird. Jedes Suchatom muß mindestens ein Wort oder einen Ausdruck enthalten.

Die folgende Anweisung ist wahr, wenn ein oder mehrere der Elemente in den Suchargumenten gefunden wird.

```
CONTAINS (mytexthandle, '( "text",
                           "graphic",
                           "audio",
                           "video")') = 1
```

PRECISE FORM OF, STEMMED FORM OF, FUZZY FORM OF, SYNONYM FORM OF, BOUND Tabelle 7 zeigt die Optionen, die den verschiedenen Typen von Indizes entsprechen. Beispielsweise ist für einen linguistischen Index jede der Optionen geeignet mit Ausnahme von PRECISE FORM OF. Wenn Sie PRECISE FORM OF angeben, wird es ignoriert und der Standardwert wird verwendet.

Die Verarbeitung von Suchbegriffen wird in Tabelle 8 auf Seite 199 näher beschrieben.

Tabelle 7. Linguistische Optionen

Schlüsselwort für Suchatom	Indextyp					
	Linguistisch	Präzise	Präzise, Normalisiert	Dual	Ngram	Ngram, Groß-/Kleinschreibung möglich
PRECISE FORM OF		X	X	X		O
STEMMED FORM OF	X			O	O	O
FUZZY FORM OF					O	O
IS ABOUT	O	O	O	O		
SYNONYM FORM OF	O	O	O	O		
EXPAND	O	O	O	O		
SOUNDS LIKE	O	O	O	O		
IN SAME SENTENCE AS	O	O	O	O	O	O
IN SAME PARAGRAPH AS	O	O	O	O	O	O
BOUND					O	O

X = Standardeinstellung, O = Funktion verfügbar

Tabelle 8. Suchbegriffsoptionen für Ngram-Indizes

Schlüsselwort für Suchatom	Verarbeitung von Suchbegriffen				
	Groß-/Kleinschreibung		Wort-stammbil-dung	Übereinstimmung	
	wird beachtet	wird nicht beachtet		Genau	Grob
PRECISE FORM OF	nur bei Beach-tung der Groß-/ Kleinschreibung	X		X	
STEMMED FORM OF		X	X		
FUZZY FORM OF		X			X
X = Standardeinstellung					

Suchargument

Wenn Sie ein Schlüsselwort verwenden, das für den betreffenden Indextyp nicht verfügbar ist, wird es ignoriert und entweder statt dessen das Standardschlüsselwort verwendet oder eine Nachricht zurückgegeben.

PRECISE FORM OF Ein Schlüsselwort, das dazu führt, daß das Wort (oder jedes Wort im Ausdruck), das auf PRECISE FORM OF folgt, genau in der eingegebenen Schreibweise gesucht wird, ohne zunächst auf die Stammform reduziert zu werden. Für präzise und duale Indizes wird bei dieser Form der Suche die Groß-/Kleinschreibung beachtet, d.&rbl;h. die Verwendung von Groß- und Kleinbuchstaben ist relevant. Wenn Sie beispielsweise nach `maus` suchen, finden Sie “Maus” nicht.

Dies ist die Standardoption für präzise und duale Indizes. Bei einem präzisen normalisierten Index wird bei der Standardform der Suche die Groß-/Kleinschreibung nicht beachtet. Wenn Sie dieses Schlüsselwort für einen linguistischen Index angeben, wird es ignoriert und STEMMED FORM OF wird angenommen.

STEMMED FORM OF Ein Schlüsselwort, das dazu führt, daß das Wort (oder jedes Wort im Ausdruck), das auf STEMMED FORM OF folgt, auf seinen Wortstamm reduziert wird, bevor die Suche ausgeführt wird. Bei dieser Form der Suche wird die Groß-/Kleinschreibung nicht beachtet. Wenn Sie beispielsweise nach `maus` suchen, finden Sie auch “Maus”.

Die Art und Weise, in der Wörter auf ihren Wortstamm reduziert werden, hängt von der Sprache ab.

Beispiel: `programming computer systems` wird bei Verwendung des Wörterverzeichnisses in amerikanischem Englisch durch `program compute system` und bei Verwendung des Wörterverzeichnisses in britischem Englisch durch `programme compute system` ersetzt.

Mit diesem Suchausdruck kann “programmer computes system”, “program computing systems”, “programming computer system” usw. gefunden werden.

Dies ist die Standardoption für linguistische Indizes. Wenn Sie dieses Schlüsselwort für einen präzisen Index angeben, wird es ignoriert und statt dessen PRECISE FORM OF angenommen.

FUZZY FORM OF Ein Schlüsselwort zum Ausführen einer “Suche nach grober Übereinstimmung”, d.&rbl;h. einer Suche nach Begriffen, die ähnlich wie der Suchbegriff geschrieben werden. Dies ist besonders nützlich, wenn in Dokumenten gesucht wird, die durch ein Programm zur optische Zeichenerkennung (OCR) erstellt wurden. Solche Dokumente enthalten oft Wörter mit Rechtschreibfehlern. Beispielsweise könnte das Wort `economy` von einem OCR-Programm als `econony` erkannt werden.

übereinstimmungsebene: Eine ganze Zahl von 1 bis 5, die den Grad der Ähnlichkeit angibt, wobei 5 eine größere Ähnlichkeit angibt als 1.

SYNONYM FORM OF Ein Schlüsselwort, das dazu führt, daß ein Wort oder ein Ausdruck, das auf SYNONYM FORM OF folgt, zusammen mit seinen Synonymen gesucht werden kann. Die Synonyme werden von dem Wörterverzeichnis zur Verfügung gestellt, das durch *sprache* angegeben ist, oder andernfalls von dem Standardwörterverzeichnis.

Synonyme für einen Ausdruck sind alternative Ausdrücke, die alle möglichen Kombinationen von Synonymen enthalten, die erhalten werden können, wenn jedes Wort des ursprünglichen Ausdrucks durch eines seiner Synonyme ersetzt wird. Die Wortreihenfolge des ursprünglichen Ausdrucks bleibt erhalten.

Wenn Sie dieses Schlüsselwort für einen präzisen Index angeben, wird es ignoriert und statt dessen PRECISE FORM OF angenommen.

Wenn Sie dieses Schlüsselwort für einen dualen Index angeben, wird die Suche unter Verwendung des linguistischen Teils des dualen Indexes durchgeführt und nicht des präzisen Teils.

Sie können dieses Schlüsselwort nicht verwenden, wenn NOT im Suchfaktor verwendet wird oder wenn das zu suchende Wort oder der zu suchende Ausdruck Platzhalterzeichen enthält.

BOUND Ein Schlüsselwort für die Suche in Dokumenten, die die koreanische CCSID verwenden. Es führt dazu, daß bei der Suche Wort- und Ausdrucksgrenzen berücksichtigt werden. Wenn *sprache* angegeben ist, wird dies ignoriert und Koreanisch wird angenommen.

Suchargument

sprache Eine Variable, die festlegt, welches Wörterverzeichnis bei der Verarbeitung auf linguistischer Basis von Textdokumenten während der Indexierung und Abfrage verwendet wird. Dies gilt nicht nur für linguistische und duale Indizes, sondern auch für präzise Indizes, da sie ein Wörterverzeichnis zur Verarbeitung von Stoppwörtern verwenden.

Zur Verarbeitung auf linguistischer Basis gehört die Synonymverarbeitung und die Wortstammverarbeitung. Weitere Informationen finden Sie in „Linguistische Funktionen für die unterstützten Sprachen“ auf Seite 30.

Die unterstützten Sprachen werden in „Sprachen“ auf Seite 258 aufgelistet.

Anmerkung: Beim Suchen in Dokumenten, die nicht in amerikanischem Englisch geschrieben sind, müssen Sie die Sprache im Suchargument angeben, *unabhängig von der Standardsprache*.

"wort oder ausdruck" Ein Wort oder ein Ausdruck, nach dem gesucht werden soll. Die Zeichen, die innerhalb eines Wortes verwendet werden können, sind sprachenabhängig. Es ist außerdem sprachenabhängig, ob Wörter durch Trennzeichen voneinander getrennt werden müssen. Für Englisch und die meisten anderen Sprachen muß jedes Wort in einem Ausdruck durch ein Leerzeichen abgetrennt sein.

Präzise oder linguistische Suche. Der Text Extender kann entweder die präzise Form eines Wortes oder Ausdrucks suchen oder eine Variation davon. Wenn Sie nicht eine der Optionen in Tabelle 7 auf Seite 198 angeben, werden die linguistischen Standardoptionen verwenden, gemäß dem verwendeten Indextyp.

Platzhalterzeichen. Ein Wort kann die folgenden Platzhalterzeichen enthalten:

_ **(Unterstreichungszeichen)** Repräsentiert ein Einzelzeichen.

% (Prozentzeichen) Repräsentiert eine beliebige Anzahl von willkürlichen Zeichen. Wenn ein Wort aus einem einzelnen % besteht, repräsentiert es ein wahlfreies Wort beliebiger Länge.

Ein Wort kann nicht ausschließlich aus Platzhalterzeichen bestehen, mit der Ausnahme, wenn ein einzelnes % verwendet wird, um ein wahlfreies Wort zu repräsentieren.

Wenn Sie ein Platzhalterzeichen verwenden, können Sie SYNONYM FORM OF, *merkmal* oder THESAURUS nicht verwenden.

ESCAPE *escape-zeichen* Ein Zeichen, das das nächste Zeichen als das zu suchende Zeichen identifiziert und nicht als ein als Platzhalterzeichen zu verwendendes Zeichen.

Beispiel: Wenn als *escape-zeichen* das Zeichen $ angegeben ist, repräsentieren %, $_ und $$ entsprechend die Zeichen %, _ und $. Alle %- und _-Zeichen, denen nicht das Zeichen $ voransteht, repräsentieren Platzhalterzeichen.

Zusammenfassung der Regeln und Einschränkungen

Boolesche Operationen NOT ist nach OR nicht zulässig.

Dualer Index Nimmt als Standardwert STEMMED FORM OF. Wenn Platzhalterzeichen verwendet werden, wird bei Suchen die Groß-/Kleinschreibung beachtet.

FUZZY FORM OF Die ersten 3 Zeichen müssen übereinstimmen. Es kann nicht verwendet werden, wenn ein Wort in dem Suchatom ein Platzhalterzeichen enthält. Es kann nicht in Kombination mit NOT verwendet werden. Es kann nur mit einem Ngram-Index verwendet werden.

IN SAME PARAGRAPH AS Es kann nicht verwendet werden, wenn NOT in einem Suchfaktor verwendet wird.

IN SAME SENTENCE AS Es kann nicht verwendet werden, wenn NOT in einem Suchfaktor verwendet wird.

Linguistischer Index Verhindert die Verwendung von PRECISE FORM OF. Nimmt als Standardwert STEMMED FORM OF. Platzhalterzeichen können verwendet werden. Bei Suchen wird die Groß-/Kleinschreibung nicht beachtet.

Platzhalterzeichen Verhindert die Verwendung von SYNONYM FORM OF, *merkmal* und THESAURUS.

Ngram-Index Platzhalterzeichen können verwendet werden, wobei allerdings kein nicht-alphanumerischen Zeichen darauf folgen kann. Bei Suchen wird die Groß-/Kleinschreibung nicht beachtet, es sei denn, der Index ist für die Beachtung der Groß-/Kleinschreibung definiert und PRECISE FORM OF wird verwendet.

Suchargument

NOT Verhindert die Verwendung von SYNONYM FORM OF, IN SAME PARAGRAPH AS und IN SAME SENTENCE AS.

PRECISE FORM OF Wird bei einem linguistischen Index ignoriert.

Präziser Index Verhindert die Verwendung von STEMMED FORM OF und SYNONYM FORM OF. Nimmt als Standardwert PRECISE FORM OF. Platzhalterzeichen können verwendet werden. Bei Suchen wird die Groß-/Kleinschreibung beachtet.

STEMMED FORM OF Wird bei einem präzisen Index ignoriert.

SYNONYM FORM OF Es kann nicht verwendet werden, wenn ein Wort in dem Suchatom ein Platzhalterzeichen enthält. Es kann nicht in Kombination mit NOT verwendet werden. Es kann nicht mit einem präzisen Index verwendet werden.

Kapitel 11. API-Funktionen zum Suchen und Anzeigen

Der Text Extender bietet C-Funktionen für das Durchsuchen von Textdokumenten und das Anzeigen der gefundenen Dokumente. Diese Funktionen bilden die Anwendungsprogrammierschnittstelle (API) des Text Extender. In diesem Kapitel werden die API-Funktionen in alphabetischer Reihenfolge beschrieben.

Kapitel 6, „Die API-Funktionen zum Suchen und Anzeigen verwenden“ auf Seite 117, liefert eine Einführung in die Funktionen und beschreibt, wie sie zusammen verwendet werden können.

Funktion	Zweck	Seite
DesCloseDocument	Gibt den Speicher frei, der von DesOpenDocument zugeordnet wurde.	206
DesEndBrowseSession	Schließt eine Anzeigesitzung und gibt den Speicher frei, der von DesStartBrowseSession zugeordnet wurde.	207
DesFreeBrowseInfo	Gibt den Speicher frei, der von DesGetBrowseInfo zugeordnet wurde.	208
DesGetBrowseInfo	Sucht unter Verwendung eines Sucharguments im Dokument nach Text und erstellt Anzeigeinformationen.	208
DesGetMatches	Gibt einen Zeiger auf Hervorhebungsinformationen für das Textdokument zurück, das durch eine Dokumentkennung beschrieben ist. Die Hervorhebungsinformationen sind ein Datenstrom. Dazu gehören der Textkontext (mindestens ein Absatz) und Informationen zum Hervorheben von Text in diesem Kontext.	211
DesGetSearchResultTable	Empfängt ein Suchargument für die Suche in Textdokumenten in einer bestimmten Textspalte und speichert das Ergebnis in einer vom Benutzer gestellten Tabelle. Sie kann außerdem Anzeigeinformationen zurückgeben.	216
DesOpenDocument	Empfängt einen Anzeigesitzungszeiger, eine Kennung und die Option DES_FAST oder DES_EXTENDED, die die Art der Verarbeitung auf linguistischer Basis angibt, die für die Hervorhebung der gefundenen Begriffe verwendet werden soll. Sie bereitet das Textdokument vor, das der Kennung entspricht, um den Dokumenttext und die Hervorhebungsinformationen abzurufen, und gibt eine Dokumentkennung zurück, die für den iterativen Aufruf von DesGetMatches verwendet wird.	220
DesStartBrowseSession	Öffnet eine Anzeigesitzung unter Verwendung der Anzeigeinformationen von DesGetBrowseInfo und gibt eine Anzeigesitzungskennung für die Verwendung durch andere Anzeigefunktionen zurück.	222

DesCloseDocument (API-Funktion)

Hinweis

Viele der API-Funktionen benötigen eine Verbindungskennung (hdbc). Sie müssen diese Kennung unter Verwendung der Funktion SQLConnect liefern, was allerdings nicht verhindert, daß Sie den Text Extender von Programmen in eingebettetem SQL aus aufrufen können. Im Handbuch *DB2 Call Level Interface Guide and Reference* wird beschrieben, wie CLI-Anweisungen und Anweisungen in eingebettetem SQL gemischt werden können.

DesCloseDocument

Zweck

Schließt ein Textdokument, das von DesOpenDocument geöffnet wurde, und gibt den Speicher, der während der Rückgabe des Dokumenttextes und der Hervorhebungsinformationen zugeordnet wurde.

Syntax

```
DESRETURN
  DesCloseDocument
     (DESBROWSESESSION     BrowseSession,
      DESHANDLE            DocumentHandle);
```

Funktionsargumente

Tabelle 9. Argumente für DesCloseDocument

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	*BrowseSession*	Eingabe	Anzeigesitzungskennung.
DESHANDLE	*DocumentHandle*	Eingabe	Kennung, die von DesOpenDocument zurückgegeben wird und die ein geöffnetes Textdokument identifiziert.

Rückkehrcodes

RC_SUCCESS

RC_INVALID_PARAMETER
RC_INVALID_SESSION
RC_SE_INCORRECT_HANDLE
RC_SE_IO_PROBLEM
RC_SE_LS_FUNCTION_FAILED
RC_SE_NOT_ENOUGH_MEMORY
RC_SE_REQUEST_IN_PROGRESS
RC_SE_WRITE_TO_DISK_ERROR

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesOpenDocument ein Textdokument geöffnet haben.

DesEndBrowseSession

Zweck

Beendet eine Anzeigesitzung, die von DesStartBrowseSession gestartet wurde, und gibt den Speicher frei, der für die Anzeigesitzung zugeordnet wurde.

Syntax

```
DESRETURN
  DesEndBrowseSession
     (DESBROWSESESSION     BrowseSession);
```

Funktionsargumente

Tabelle 10. Argumente für DesEndBrowseSession

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	*BrowseSession*	Eingabe	Anzeigesitzungskennung.

Verwendung

Diese Funktion gibt den Speicher, der durch DesGetBrowseInfo für die Anzeigesitzung zugeordnet wurde, nicht frei. Dieser Speicher enthält Anzeigeinformationen, die noch von einer anderen Anzeigesitzung verwendet werden könnten. Um diesen Speicher freizugeben, rufen Sie DesFreeBrowseInfo auf.

Rückkehrcodes

RC_SUCCESS

RC_INVALID_SESSION
RC_INVALID_PARAMETER
RC_SE_UNEXPECTED_ERROR

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesStartBrowseSession eine Anzeigesitzung gestartet haben.

DesGetBrowseInfo (API-Funktion)

DesFreeBrowseInfo

Zweck

Gibt den Speicher frei, der durch DesGetBrowseInfo für die Anzeigeinformationen zugeordnet wurde.

Syntax

```
DESRETURN
  DesFreeBrowseInfo
     (DESBROWSEINFO        BrowseInfo );
```

Funktionsargumente

Tabelle 11. Argumente für DesFreeBrowseInfo

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSEINFO	*BrowseInfo*	Eingabe	Anzeigeinformationen.

Rückkehrcodes

RC_SUCCESS

RC_INVALID_PARAMETER

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesGetBrowseInfo Speicher für die Anzeigeinformationen zugeordnet haben.

DesGetBrowseInfo

Zweck

Empfängt ein Suchargument für das Durchsuchen von Text, der durch eine Kennung identifiziert ist. Die Funktion gibt einen Zeiger auf Anzeigeinformationen zurück, die von DesStartBrowseSession zum Hervorheben der gefundenen Begriffe benötigt wird.

DesGetBrowseInfo (API-Funktion)

Syntax

```
DESRETURN
  DesGetBrowseInfo
   (SQLHDBC             hdbc,
    SQLCHAR             *pHandle,
    DESUSHORT           HandleLength,
    char                *pSearchArgument,
    DESSMALLINT         ArgumentLength,
    DESBROWSEINFO       *pBrowseInfo,
    DESMESSAGE          *pErrorMessage);
```

Funktionsargumente

Tabelle 12. Argumente für DesGetBrowseInfo

Datentyp	Argument	Verwendung	Beschreibung
SQLHDBC	*hdbc*	Eingabe	Eine Datenbankverbindungskennung.
SQLCHAR *	*pHandle*	Eingabe	Zeiger auf eine Kennung, die aus der Datenbank extrahiert wurde.
DESUSHORT	*HandleLength*	Eingabe	Länge von pHandle. DES_NTS kann hier nicht verwendet werden.
char *	*pSearchArgument*	Eingabe	Zeiger auf ein Textsuchargument, das die Informationen angibt, die Sie suchen wollen.
DESSMALLINT	*ArgumentLength*	Eingabe	Entweder die Länge von pSearchArgument (ohne Nullbyteabschlußzeichen) oder DES_NTS.
DESBROWSEINFO *	*pBrowseInfo*	Ausgabe	Zeiger auf Anzeigeinformationen, die die Daten enthalten, die zum Anzeigen eines Textdokuments erforderlich sind. Dieser Zeiger wird an DesStartBrowseSession übergeben.
DESMESSAGE *	*pErrorMessage*	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der Text Extender einen Fehlercode und eine Fehlernachricht zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn *pErrorMessage* der Nullzeiger ist, wird keine Fehlernachricht zurückgegeben.

DesGetBrowseInfo (API-Funktion)

Verwendung

Ihr Anwendungsprogramm muß eine Verbindung zur Datenbank herstellen, bevor es DesGetBrowseInfo aufruft.

Als Zeiger auf das Suchargument wird eher *char** anstelle von *SQLCHAR** verwendet. Der Grund dafür ist, daß der Parameterwert möglicherweise nicht aus der Datenbank stammt.

Für die Zuordnung zwischen SQL-Datentypen und C-Datentypen müssen Sie den symbolischen SQL-Namen SQL_VARBINARY für eine Kennung verwenden. Der Typ der Host-Variablen, die auf die C-Repräsentation von *Handle*-Werten zeigen, ist SQLCHAR*.

Der Text Extender ordnet Speicher für Anzeigeinformationen zu. Das Anwendungsprogramm muß den Speicher und zugehörige Ressourcen durch Aufrufen von DesFreeBrowseInfo freigeben.

Da *Handle*-Werte Bitdaten sind und mehrere '\0'-Zeichen enthalten, müssen Sie die Länge von *pHandle* angeben.

Das Suchargument in pSearchArgument wird in Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, beschrieben.

Rückkehrcodes

RC_SUCCESS
RC_NO_BROWSE_INFO

RC_ALLOCATION_ERROR
RC_FILE_IO_PROBLEM
RC_INTERNAL_ERROR
RC_INVALID_PARAMETER
RC_PARSER_INVALID_ESCAPE_CHARACTER
RC_PARSER_INVALID_USE_OF_ESCAPE_CHAR
RC_PARSER_SYNTAX_ERROR
RC_SE_COMMUNICATION_PROBLEM
RC_SE_EMPTY_INDEX
RC_SE_EMPTY_QUERY
RC_SE_FUNCTION_DISABLED
RC_SE_FUNCTION_IN_ERROR
RC_SE_INCORRECT_HANDLE
RC_SE_INDEX_DELETED
RC_SE_INDEX_NOT_ACCESSIBLE
RC_SE_INDEX_SUSPENDED
RC_SE_INSTALLATION_PROBLEM
RC_SE_IO_PROBLEM
RC_SE_MAX_NUMBER_OF_BUSY_INDEXES
RC_SE_MAX_OUTPUT_SIZE_EXCEEDED
RC_SE_NOT_ENOUGH_MEMORY

RC_SE_PROCESSING_LIMIT_EXCEEDED
RC_SE_QUERY_TOO_COMPLEX
RC_SE_SERVER_BUSY
RC_SE_SERVER_CONNECTION_LOST
RC_SE_SERVER_NOT_AVAILABLE
RC_SE_UNEXPECTED_ERROR
RC_SE_UNKNOWN_INDEX_NAME
RC_SE_UNKNOWN_SERVER_NAME
RC_SE_WRITE_TO_DISK_ERROR

Achtung: Die nachfolgenden Rückkehrcodes geben an, daß die Funktion ein Ergebnis, aber möglicherweise nicht das erwartete zurückgegeben hat.

RC_SE_CONFLICT_WITH_INDEX_TYPE
RC_SE_DICTIONARY_NOT_FOUND
RC_SE_STOPWORD_IGNORED
! RC_SE_UNKNOWN_SECTION_NAME
! RC_SE_DOCMOD_READ_PROBLEM

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie eine Verbindung zu einer Datenbank hergestellt haben und eine Text Extender-UDF verwendet haben, um eine Kennung aus der Datenbank zu extrahieren.

DesGetMatches

Zweck

Gibt einen Datenstrom zurück, der Hervorhebungsinformationen für das Textdokument enthält, das durch eine Dokumentkennung beschrieben ist. Siehe „Datenstromsyntax“ auf Seite 213. Zu Hervorhebungsinformationen gehören der Textkontext (mindestens ein Absatz) und Informationen zum Hervorheben von Text in diesem Kontext.

DesGetMatches gibt nur einen Abschnitt des Datenstroms zurück, wobei die Länge des Abschnitts in der Ausgabestruktur angegeben wird.

Eine Folge von Aufrufen von DesGetMatches ruft den Inhalt des gesamten Textdokuments ab. Wenn das Ende des Textdokuments erreicht ist, wird RC_SE_END_OF_INFORMATION zurückgegeben.

DesGetMatches (API-Funktion)

Syntax

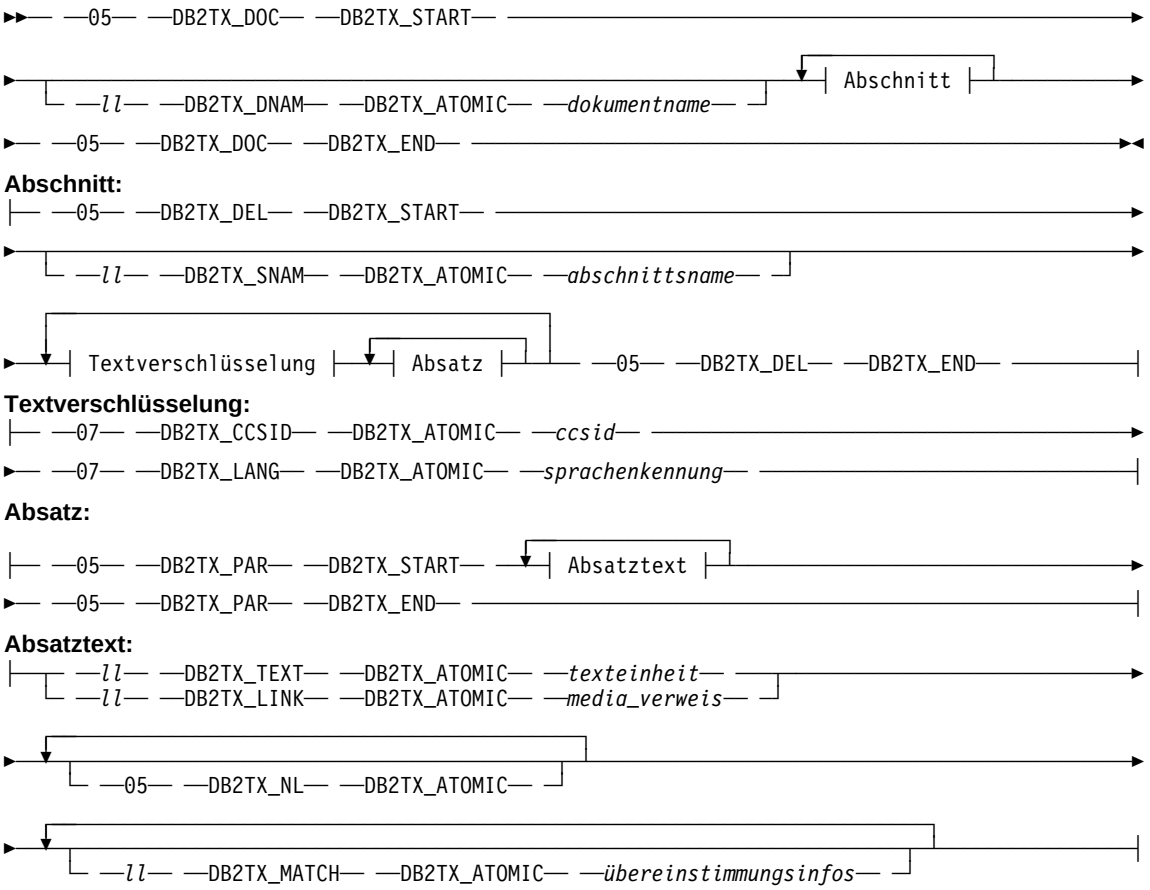
```
DESRETURN
DesGetMatches
    (DESBROWSESESSION    BrowseSession,
      DESHANDLE           DocumentHandle,
      DESMATCHINFO        *pMatchInfo,
      DESULONG            *pMatchInfoLength,
      DESMESSAGE          *pErrorMessage);
```

Funktionsargumente

Tabelle 13. Argumente für DesGetMatches

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	*BrowseSession*	Eingabe	Anzeigesitzungskennung.
DESHANDLE	*DocumentHandle*	Eingabe	Dokumentkennung, die von DesOpenDocument zurückgegeben wird.
DESMATCHINFO *	*pMatchInfo*	Ausgabe	Zeiger auf einen Puffer, der die empfangenen Datenstromabschnitte enthält. DesGetMatches ordnet diesen Puffer zu.
DESULONG *	*pMatchInfoLength*	Ausgabe	Die Länge des Datenstromabschnitts, auf den durch pMatchInfo gezeigt wird.
DESMESSAGE *	*pErrorMessage*	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der Text Extender einen Fehlercode und eine Fehlernachricht zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn *pErrorMessage* der Nullzeiger ist, wird keine Fehlernachricht zurückgegeben.

Datenstromsyntax



Jedes Segment im Syntaxdiagramm, wie z.&rbl;B. `05 DB2TX_DOC DB2TX_START` beginnt mit einem Längenfeld vom Typ Integer, das in diesem Diagramm entweder eine explizite Zahl, z.&rbl;B. `05`, oder eine Variable *ll* ist. Die Länge des Segments schließt das aus 2 Byte bestehende Längenfeld ein.

Anmerkung: Die Länge ist im Big-Endian-Format angegeben.

Zu jedem Segment gehört eine der folgenden, aus einem Byte bestehende Typenkennungen:

DB2TX_START
Gibt den Start eines Segments an, z.&rbl;B. Dokument oder Absatz.

DB2TX_END
Gibt das Ende eines Segments an.

DesGetMatches (API-Funktion)

DB2TX_ATOMIC
Gibt an, daß das nachfolgende Element autark ist, z.&rbl;B. ein Dokumentname oder eine Sprachenkennung.

Die Datenstromelemente sind jeweils 2 Byte lang. Dazu gehören:

DB2TX_DOC
Gibt den Start und das Ende eines Dokuments an.

DB2TX_DNAM
Ein Dokumentname. Wenn kein Name angegeben ist, wird die Kennung des Dokuments verwendet.

DB2TX_DEL
Gibt den Start und das Ende eines Dokumentelements an. Der einzige Typ von Dokumentelement, der momentan unterstützt wird, ist ein *Textabschnitt*.

DB2TX_SNAM
Gibt den Namen eines Textabschnitts an. Momentan unterstützt der Text Extender nur einen Textabschnitt und liefert automatisch einen Standardnamen. Wenn Sie einen Abschnittsnamen angeben, wird er ignoriert.

DB2TX_PAR
Gibt den Start und das Ende eines Textabsatzes innerhalb des aktuellen Abschnitts an.

DB2TX_TEXT
Gibt einen Textbereich innerhalb des aktuellen Absatzes an. Normalerweise enthält die *Texteinheit* eine Zeile mit Text, und auf das Element TEXT folgt ein Element DB2TX_NL; allerdings können Textzeilen möglicherweise auch in mehrere Teile aufgeteilt sein, wobei jeder Teil in seinem eigenen Element DB2TX_TEXT angegeben ist.

Der Text verwendet die CCSID und Sprache, die dem aktuellen Absatz zugeordnet sind.

DB2TX_LINK
Gibt einen Text Extender-Hypermedia-Verweis an. Die CCSID des aktuellen Absatzes wird verwendet.

DB2TX_NL
Gibt den Start einer neuen Zeile im aktuellen Absatz an.

DB2TX_MATCH
Enthält Informationen zum Vorkommen von Übereinstimmungen im aktuellen Textabschnitt. Die Informationen werden als Folge von Binärzahlpaaren geliefert. Die erste Zahl in jedem Paar ist die relative Position einer Übereinstimmung innerhalb des aktuellen Textabschnitts, die zweite Zahl ist die Länge (in Zeichen) dieser Übereinstimmung. Die angegebene Länge kann den betroffenen Textabschnitt überschreiten. Sowohl relative Position als auch Länge sind aus zwei Byte bestehende Werte, die im Big-Endian-Format angegeben sind.

DB2TX_CCSID
Die CCSID für Text in nachfolgenden Absätzen, bis einem Absatz ein neues Element DB2TX_CCSID vorausgeht. Die folgenden CCSIDs werden zurückgegeben:

`DB2TX_CCSID_00500`
für Text in der EBCDIC-Codepage 500 für das lateinische Alphabet 1.

`DB2TX_CCSID_04946`
für Text in der ASCII-Codepage 850 für das lateinische Alphabet 1.

`DB2TX_CCSID_00819`
für Text in der ASCII-Codepage 819.

Diese symbolischen Namen für CCSIDs sind in der Datei DES_EXT.H definiert, die mit dem Text Extender geliefert wird. Die aus zwei Byte bestehenden Werte sind im Big-Endian-Format angegeben.

DB2TX_LANG
Die Sprachenkennung für Text in nachfolgenden Absätzen, bis einem Absatz ein neues Element DB2TX_LANG vorausgeht. Die Datei DES_EXT.H, die mit dem Text Extender geliefert wird, definiert symbolische Namen für alle Sprachenkennungen, die vom Text Extender unterstützt werden. Die aus zwei Byte bestehenden Werte sind im Big-Endian-Format angegeben.

Verwendung

DesGetMatches gibt RC_SE_END_OF_INFORMATION zurück, wenn das Ende des Textdokuments erreicht ist.

Rückkehrcodes

RC_SUCCESS
RC_SE_END_OF_INFORMATION

RC_INVALID_PARAMETER
RC_INVALID_SESSION
RC_SE_CAPACITY_LIMIT_EXCEEDED
RC_SE_INCORRECT_HANDLE
RC_SE_IO_PROBLEM
RC_SE_NOT_ENOUGH_MEMORY
RC_SE_REQUEST_IN_PROGRESS
RC_SE_LS_FUNCTION_FAILED
RC_SE_UNEXPECTED_ERROR

Achtung: Die nachfolgenden Rückkehrcodes geben an, daß die Funktion ein Ergebnis, aber möglicherweise nicht das erwartete zurückgegeben hat.

RC_SE_DICTIONARY_NOT_FOUND

DesGetSearchResultTable (API-Funktion)

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesOpenDocument ein Textdokument geöffnet haben.

DesGetSearchResultTable

Zweck

Verwendet ein Suchargument für das Durchsuchen von Textdokumenten, die durch eine Textspalte identifiziert sind. Die Kennungsdaten der gefundenen Textelemente werden in eine Ergebnistabelle geschrieben. Anzeigeinformationen zur Rangordnung und der Anzahl der Übereinstimmungen können ebenfalls in die Ergebnistabelle geschrieben werden.

Syntax

```
DESRETURN
  DesGetSearchResultTable
    (SQLHDBC           hdbc,
     char              *pTableSchema,
     DESSMALLINT       TableSchemaLength,
     char              *pTableName,
     DESSMALLINT       TableNameLength,
     char              *pColumnName,
     DESSMALLINT       ColumnNameLength,
     char              *pSearchArgument,
     DESSMALLINT       ArgumentLength,
     char              *pResultSchema,
     DESSMALLINT       ResultSchemaLength,
     char              *pResultTableName,
     DESSMALLINT       ResultTableNameLength,
     DESSEARCHOPTION   SearchOption,
     DESBROWSEOPTION   BrowseOption,
     DESBROWSEINFO     *pBrowseInfo,
     DESMESSAGE        *pErrorMessage);
```

Funktionsargumente

Tabelle 14 (Teil 1 von 3). Argumente für DesGetSearchResultTable

Datentyp	Argument	Verwendung	Beschreibung
SQLHDBC	hdbc	Eingabe	Eine Datenbankverbindungskennung.
char *	pTableSchema	Eingabe	Das Schema der zu durchsuchenden Basistabelle.
DESSMALLINT	TableSchemaLength	Eingabe	Entweder die Länge von pTableSchema (ohne Nullbyteabschlußzeichen) oder DES_NTS.

Tabelle 14 (Teil 2 von 3). Argumente für DesGetSearchResultTable

Datentyp	Argument	Verwendung	Beschreibung
char *	pTableName	Eingabe	Zeiger auf den Namen der zu durchsuchenden Basistabelle.
DESSMALLINT	TableNameLength	Eingabe	Entweder die Länge von pTableName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pColumnName	Eingabe	Zeiger auf den Namen der Spalte, die in der beabsichtigten Textsuche adressiert werden soll. Die Spalte muß vom Typ DESTEXTH sein.
DESSMALLINT	ColumnNameLength	Eingabe	Entweder die Länge von pColumnName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pSearchArgument	Eingabe	Zeiger auf das Textsuchargument.
DESSMALLINT	ArgumentLength	Eingabe	Entweder die Länge von pSearchArgument (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pResultSchema	Eingabe	Zeiger auf das Schema, das die Ergebnistabelle enthält.
DESSMALLINT	ResultSchemaLength	Eingabe	Entweder die Länge von pSchemaName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
char *	pResultTableName	Eingabe	Zeiger auf den Namen der Ergebnistabelle, die Sie zuvor erstellt haben, in der das Ergebnis der Suche gespeichert werden soll. Die Struktur dieser Tabelle finden Sie in Abb. 16 auf Seite 218.
DESSMALLINT	ResultTableNameLength	Eingabe	Entweder die Länge von pResultTableName (ohne Nullbyteabschlußzeichen) oder DES_NTS.
DESSEARCHOPTION	SearchOption	Eingabe	Eine Option, die festlegt, ob die Rangordnungsinformationen, die Anzahl an Übereinstimmungen oder nur die Kennungen der übereinstimmenden Textdokumente abgerufen werden. DES_RANK DES_MATCH DES_RANKANDMATCH DES_TEXTHANDLEONLY Diese Option legt den Inhalt der Ergebnistabelle fest, wie in „Verwendung“ auf Seite 218 beschrieben.
DESBROWSEOPTION	BrowseOption	Eingabe	Reserviert.

DesGetSearchResultTable (API-Funktion)

Tabelle 14 (Teil 3 von 3). Argumente für DesGetSearchResultTable

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSEINFO *	pBrowseInfo	Ausgabe	Zeiger auf Anzeigeinformationen oder ein Zeiger auf Null, je nachdem, welchen Wert BrowseOption hat.
DESMESSAGE *	*pErrorMessage*	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der Text Extender einen Fehlercode und eine Fehlernachricht zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn *pErrorMessage* der Nullzeiger ist, wird keine Fehlernachricht zurückgegeben.

Verwendung

Die Verbindung zur Datenbank muß vom Anwendungsprogramm durch Aufrufen von DesGetSearchResultTable hergestellt worden sein.

Der Name *pResultTableName* bezieht sich auf eine Ergebnistabelle, die Sie im voraus erstellt haben. Das Dienstprogramm DESRESTB im Beispielverzeichnis erstellt eine Ergebnistabelle für Textkennungen. Nach dem Aufruf dieser Funktion enthält die Ergebnistabelle Informationen, die Textwerte identifizieren, die mit dem Suchargument übereinstimmen. Dies ist die Struktur der Ergebnistabelle:

ERGEBNISTABELLE

TEXTHANDLE	RANK	MATCHES

Abbildung 16. Struktur der Ergebnistabelle

Der Datentyp von TEXTHANDLE ist DB2TEXTH oder DB2TEXTFH. Der Datentyp von RANK ist DOUBLE. Der Datentyp von MATCHES ist INTEGER.

Das Suchargument in pSearchArgument wird in Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, beschrieben.

Wenn der Wert von *BrowseOption BROWSE* ist, gibt der Text Extender Anzeigeinformationen aus der Text Extender-Suchmaschine zurück, die sich auf dem Server befindet. *pBrowseInfo* zeigt auf die Anzeigeinformationen, die als Eingabe für DesStartBrowseSession verwendet werden. Wenn der Wert von *BrowseOption NO_BROWSE* ist, zeigt *pBrowseInfo* auf Null.

Rückkehrcodes

RC_SUCCESS
RC_NO_BROWSE_INFO
RC_SE_NO_DATA
RC_ALLOCATION_ERROR
RC_FILE_IO_PROBLEM
RC_INTERNAL_ERROR
RC_INVALID_BROWSE_OPTION
RC_INVALID_PARAMETER
RC_INVALID_SEARCH_OPTION
RC_INVALID_SESSION
RC_PARSER_INVALID_ESCAPE_CHARACTER
RC_PARSER_SYNTAX_ERROR
RC_RESULT_TABLE_NOT_EXIST
RC_SE_COMMUNICATION_PROBLEM
RC_SE_EMPTY_INDEX
RC_SE_EMPTY_QUERY
RC_SE_FUNCTION_DISABLED
RC_SE_FUNCTION_IN_ERROR
RC_SE_INCORRECT_HANDLE
RC_SE_INDEX_DELETED
RC_SE_INDEX_NOT_ACCESSIBLE
RC_SE_INDEX_SUSPENDED
RC_SE_INSTALLATION_PROBLEM
RC_SE_IO_PROBLEM
RC_SE_MAX_NUMBER_OF_BUSY_INDEXES
RC_SE_NOT_ENOUGH_MEMORY
RC_SE_PROCESSING_LIMIT_EXCEEDED
RC_SE_QUERY_TOO_COMPLEX
RC_SE_SERVER_BUSY
RC_SE_SERVER_CONNECTION_LOST
RC_SE_SERVER_NOT_AVAILABLE
RC_SE_UNEXPECTED_ERROR
RC_SE_UNKNOWN_INDEX_NAME
RC_SE_UNKNOWN_SERVER_NAME
RC_SE_WRITE_TO_DISK_ERROR
RC_SQL_ERROR_NO_INFO
RC_SQL_ERROR_WITH_INFO
RC_TEXT_COLUMN_NOT_ENABLED

Achtung: Die nachfolgenden Rückkehrcodes geben an, daß die Funktion ein Ergebnis, aber möglicherweise nicht das erwartete zurückgegeben hat.

RC_SE_CONFLICT_WITH_INDEX_TYPE
RC_SE_DICTIONARY_NOT_FOUND
RC_SE_STOPWORD_IGNORED
! RC_SE_UNKNOWN_SECTION_NAME
! RC_SE_DOCMOD_READ_PROBLEM

DesOpenDocument (API-Funktion)

DesOpenDocument

Zweck

Empfängt einen Anzeigesitzungszeiger, eine Kennung und eine Option (DES_EXTENDED oder DES_FAST), die angibt, ob das Textdokument mit oder ohne Verwendung eines Wörterverzeichnisses analysiert werden soll. Diese Funktion bereitet das Textdokument vor, das der Kennung entspricht, um den Dokumenttext und die Hervorhebungsinformationen abzurufen, und gibt eine Dokumentkennung zurück, die für den iterativen Aufruf von DesGetMatches verwendet wird.

Syntax

```
DESRETURN
  DesOpenDocument
    (DESBROWSESESSION    BrowseSession,
     SQLCHAR             *pHandle,
     DESUSHORT           HandleLength,
     DESMATCHMODE        MatchMode,
     DESHANDLE           *pDocumentHandle,
     DESMESSAGE          *pErrorMessage);
```

Funktionsargumente

Tabelle 15 (Teil 1 von 2). Argumente für DesOpenDocument

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSESESSION	*BrowseSession*	Eingabe	Anzeigesitzungskennung.
SQLCHAR *	*pHandle*	Eingabe	Zeiger auf eine Kennung, die aus der Datenbank extrahiert wurde.
DESUSHORT	HandleLength	Eingabe	Länge von pHandle (DES_NTS kann nicht verwendet werden).
DESMATCHMODE	MatchMode	Eingabe	Modus zum Ermitteln, ob ein Wörterverzeichnis zum Suchen von Hervorhebungsinformationen verwendet wird. **DES_FAST** Kein Wörterverzeichnis verwenden **DES_EXTENDED** Ein Wörterverzeichnis verwenden
DESHANDLE *	pDocumentHandle	Ausgabe	Eine Dokumentkennung für das iterative Aufrufen von DesGetMatches.

Tabelle 15 (Teil 2 von 2). Argumente für DesOpenDocument

Datentyp	Argument	Verwendung	Beschreibung
DESMESSAGE *	*pErrorMessage*	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der Text Extender einen Fehlercode und eine Fehlernachricht zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn *pErrorMessage* der Nullzeiger ist, wird keine Fehlernachricht zurückgegeben.

Verwendung

DES_FAST und DES_EXTENDED beziehen sich auf die Verwendung der Verarbeitung auf linguistischer Basis für die Ermittlung, welche Begriffe in dem angezeigten Text hervorgehoben werden sollen. Weitere Informationen finden Sie in „Verarbeitung auf linguistischer Basis für die Anzeige“ auf Seite 27. Geben Sie DES_FAST an, um die Basistextanalyse zu verwenden, bzw. DES_EXTENDED, um die erweiterte Übereinstimmung zu verwenden.

Für die Zuordnung zwischen SQL-Datentypen und C-Datentypen müssen Sie den symbolischen SQL-Namen SQL_VARBINARY für eine Kennung verwenden. Der Typ der Host-Variablen, die auf die C-Repräsentation von *TextHandle*-Werten zeigen, ist SQLCHAR*.

Der Text Extender ordnet Speicher für Anzeigeinformationen zu. Das Anwendungsprogramm muß diesen Speicher und zugehörige Ressourcen durch Aufrufen von DesFreeBrowseInfo freigeben.

Da *TextHandle*-Werte Bitdaten sind und mehrere '\0'-Zeichen enthalten, müssen Sie die Länge von *pHandle* angeben.

Das aufrufende Programm muß Lesezugriff auf die Tabelle haben, die das Textdokument enthält, auf das pHandle verweist.

Rückkehrcodes

RC_SUCCESS

RC_ALLOCATION_ERROR
RC_INTERNAL_ERROR
RC_INVALID_MATCH_OPTION
RC_INVALID_PARAMETER
RC_INVALID_SESSION
RC_SE_DOCUMENT_NOT_ACCESSIBLE
RC_SE_DOCUMENT_NOT_FOUND
RC_SE_INCORRECT_HANDLE

DesStartBrowseSession (API-Funktion)

RC_SE_IO_PROBLEM
RC_SE_LS_FUNCTION_FAILED
RC_SE_LS_NOT_EXECUTABLE
RC_SE_MAX_NUMBER_OF_BUSY_INDEXES
RC_SE_NOT_ENOUGH_MEMORY
RC_SE_REQUEST_IN_PROGRESS
RC_SE_UNKNOWN_INDEX_NAME
RC_SE_UNEXPECTED_ERROR

Einschränkungen

Diese Funktion kann nur aufgerufen werden, nachdem Sie durch einen Aufruf von DesStartBrowseSession eine Anzeigesitzung gestartet haben.

DesStartBrowseSession

Zweck

Startet eine Anzeigesitzung, indem eine Umgebung hergestellt wird, die zum Anzeigen eines Textdokuments und zum Hervorheben von dessen Übereinstimmungen erforderlich ist. Diese Funktion empfängt einen Zeiger auf Anzeigeinformationen, entweder von DesGetBrowseInfo oder von DesGetSearchResultTable, und gibt eine Anzeigesitzungskennung für die Verwendung durch andere Anzeigefunktionen zurück.

Syntax

```
! DESRETURN
!   DesStartBrowseSession
!      (DESBROWSEINFO       BrowseInfo,
!       char                *pUserId,
!       DESSMALLINT          UserIdLength,
!       char                *pPassword,
!       DESSMALLINT          PasswordLength,
!       DESBROWSESESSION    *pBrowseSession,
!       DESMESSAGE          *pErrorMessage);
```

Funktionsargumente

Tabelle 16 (Teil 1 von 2). Argumente für DesStartBrowseSession

Datentyp	Argument	Verwendung	Beschreibung
DESBROWSEINFO	*BrowseInfo*	Eingabe	Zeiger auf Informationen, die zum Anzeigen und Hervorheben von Übereinstimmungen in einem Textdokument benötigt werden. Der Zeiger wird von DesGetSearchResultTable oder DesGetBrowseInfo zurückgegeben.
! char *	*pUserId*	Eingabe	Benutzer-ID für die Datenbank.

Tabelle 16 (Teil 2 von 2). Argumente für DesStartBrowseSession

Datentyp	Argument	Verwendung	Beschreibung
DESSMALLINT	*UserIdLength*	Eingabe	Länge der Benutzer-ID für die Datenbank.
char *	*pPassword*	Eingabe	Kennwort für die Datenbank.
DESSMALLINT	*PasswordLength*	Eingabe	Länge des Kennworts für die Datenbank.
DESBROWSESESSION *	*pBrowseSession*	Ausgabe	Eine Kennung für eine Anzeigesitzung zur Verwendung durch andere Anzeigefunktionen.
DESMESSAGE *	*pErrorMessage*	Ausgabe	Durch die Implementierung definierter Nachrichtentext. Wenn ein Fehler auftritt, gibt der Text Extender einen Fehlercode und eine Fehlernachricht zurück. Das Anwendungsprogramm ordnet den Puffer der Größe DES_MAX_MESSAGE_LENGTH zu. Wenn *pErrorMessage* der Nullzeiger ist, wird keine Fehlernachricht zurückgegeben.

Verwendung

Diese Funktion öffnet eine Anzeigesitzung zum Anzeigen von Textdokumenten. Sie werden aufgefordert, Ihre Benutzer-ID und das Kennwort einzugeben, damit Ihre Berechtigung für den Zugriff auf die Datenbank überprüft wird.

Sie schließen die Anzeigesitzung durch Aufrufen von DesEndBrowseSession.

BrowseInfo hängt vom Suchargument ab und von der Basistextspalte, die für die Erstellung der Anzeigeinformationen verwendet wurde.

Rückkehrcodes

RC_SUCCESS

RC_ALLOCATION_ERROR
RC_INVALID_BROWSE_INFO
RC_INVALID_PARAMETER
RC_INTERNAL_ERROR
RC_SE_NOT_ENOUGH_MEMORY
RC_SE_UNEXPECTED_ERROR
RC_SQL_ERROR_NO_INFO
RC_SQL_ERROR_WITH_INFO

Einschränkungen

Sie müssen DesGetBrowseInfo oder DesGetSearchResultTable mit der entsprechenden Browse-Option aufrufen, bevor Sie diese Funktion aufrufen.

DesStartBrowseSession (API-Funktion)

Kapitel 12. Rückkehrcodes

In diesem Kapitel werden die Codes aufgelistet, die von der Text Extender-API als Antwort auf einen Funktionsaufruf zurückgegeben werden. Sie sind in alphabetischer Reihenfolge aufgelistet.

Alle Text Extender-API-Aufrufe geben als C-Funktionswert einen numerischen Rückkehrcode zurück. Die Rückkehrcodes sind in der Kopfdatei DES_EXT.H definiert, die mit dem Text Extender geliefert wird.

Die Text Extender-API fängt Fehlersituationen ab und berichtet Fehlerbedingungen mit einem Rückkehrcode.

Anwendungen, die Text Extender-API-Funktionen aufrufen, sollten immer den Rückkehrcode prüfen, bevor Sie versuchen, andere Ausgabeparameter zu verarbeiten. Die für jeden Aufruf jeweils möglichen Rückkehrcodes werden mit ihren Parametern in Kapitel 11, „API-Funktionen zum Suchen und Anzeigen“ auf Seite 205, aufgelistet.

In einigen Fällen kann eine falsche Eingabe, z.&rbl;B. ein veralteter Sitzungszeiger, zu einer abnormalen Beendigungsbedingung in den API-Services führen, die nicht durch den Text Extender abgefangen werden kann.

RC_ALLOCATION_ERROR

Erläuterung: Speicher für die interne Verwendung kann nicht zugeordnet werden.

Aktion: Stellen Sie sicher, daß ausreichend Speicher verfügbar ist.

RC_FILE_IO_PROBLEM

Erläuterung: Der Text Extender konnte eine Datei nicht lesen oder schreiben.

Aktion: Überprüfen Sie, daß auf dem Server ausreichend Plattenspeicherplatz und Hauptspeicher verfügbar ist. Überprüfen Sie, daß die Umgebungsvariablen und die Textkonfigurationseinstellungen korrekt sind.

RC_INVALID_BROWSE_INFO

Erläuterung: Die Anzeigeinformationen, die von DesGetSearchResultTable oder DesGetBrowseInfo zurückgegeben wurden und als Eingabe für DesStartBrowseSession verwendet werden, sind nicht gültig.

Aktion: Überprüfen Sie, ob ein Programmierfehler die Anzeigeinformationen überschreibt.

RC_INVALID_BROWSE_OPTION

Erläuterung: Die Anzeigeoption in DesGetSearchResultTable ist nicht gültig.

Aktion: Stellen Sie sicher, daß die Option BROWSE oder NO_BROWSE ist.

RC_INVALID_MATCH_OPTION

Erläuterung: Die in DesOpenDocument verwendeten Übereinstimmungsoptionen sind nicht gültig.

Aktion: Stellen Sie sicher, daß die Option FAST oder EXTENDED ist.

RC_INVALID_PARAMETER

Erläuterung: Einer der Eingabeparameter ist falsch.

Aktion: Lesen Sie die Fehlernachricht, die vom Text Extender zurückgegeben wurde, um die Ursache festzustellen.

RC_INVALID_SEARCH_OPTION

Erläuterung: Die Suchoption in DesGetSearchResultTable ist nicht gültig.

Aktion: Stellen Sie sicher, daß die Option DES_TEXTHANDLEONLY, DES_RANK, DES_MATCH oder DES_RANKANDMATCH ist.

RC_INVALID_SESSION

Erläuterung: Der im aktuellen Serviceaufruf angegebene Sitzungszeiger ist falsch oder veraltet.

Aktion: Speichern Sie alle Informationen, die Sie bei der Suche nach der Fehlerursache unterstützen können, und beenden Sie anschließend die Anwendung.

RC_NO_BROWSE_INFO

Erläuterung: Vom Text Extender wurden keine Anzeigeinformationen zurückgegeben. Ursache dafür ist, daß das Suchargument zu einem leeren Suchergebnis führte. Dies ist kein Fehler.

Aktion: Keine Aktion erforderlich.

RC_PARSER_INVALID_ESCAPE_CHARACTER

Erläuterung: Die Suchkriterien enthalten ein falsches Escape-Zeichen. Dieser Fehler wird gemeldet, wenn ein Leerzeichen als Escape-Zeichen verwendet wird oder wenn für ein Wort oder einen Ausdruck mehr als ein Escape-Zeichen in den Suchkriterien angegeben wurde. Beispiel: `ESCAPE " "` oder `ESCAPE "#$"`.

Aktion: Überprüfen Sie die Syntax des Sucharguments und wiederholen Sie den Versuch.

RC_PARSER_INVALID_USE_OF_ESCAPE_CHAR

Erläuterung: Die Escape-Zeichensyntax in den Suchkriterien kann nicht interpretiert werden.

Aktion: Überprüfen Sie die Escape-Zeichensyntax. Wenn beispielsweise $ das angegebene Escape-Zeichen ist, kann das Wort oder der Ausdruck nur $$, $_ oder $% enthalten, wobei _ und % die zwei Platzhalterzeichen sind.

RC_PARSER_SYNTAX_ERROR

Erläuterung: Die Suchkriteriensyntax kann nicht interpretiert werden.

Aktion: Überprüfen Sie die Syntax des Sucharguments, indem Sie Kapitel 10, „Syntax von Suchargumenten“ auf Seite 193, verwenden.

RC_RESULT_TABLE_NOT_EXIST

Erläuterung: Sie versuchen, das Ergebnis einer Suche in einer Tabelle zu speichern, die nicht existiert.

Aktion: Erstellen Sie die Tabelle, wie in Abb. 16 auf Seite 218 gezeigt.

RC_SE_BROWSER_TIME_OUT

Erläuterung: Der Anzeigeprozeß wurde gestartet, jedoch erfolgte keine Antwort innerhalb der zulässigen Zeit. Der Text Extender hat anschließend den anstehenden Prozeß abgebrochen.

Dieser Fehler kann auftreten, wenn das System nicht über genügend Speicherbereich verfügt oder überlastet ist.

Aktion: Beenden Sie die Anzeigesitzung, indem Sie DesEndBrowseSession aufrufen, geben Sie durch Aufrufen von DesFreeBrowseInfo Speicher frei und wiederholen Sie den Versuch.

RC_SE_CAPACITY_LIMIT_EXCEEDED

Erläuterung: Die angeforderte Funktion kann nicht verarbeitet werden. Nicht genügend Speicher oder Plattenspeicherplatz.

Aktion: Beenden Sie das Programm und überprüfen Sie die Ressourcen Ihres Systems.

RC_SE_COMMUNICATION_PROBLEM

Erläuterung: Die Kommunikation mit dem Text Extender-Server ist fehlgeschlagen. Der Fehler wurde möglicherweise durch fehlenden Speicherbereich oder durch eine inkorrekte Installation des Text Extender verursacht.

Aktion: Speichern Sie alle Informationen, die Sie bei der Suche nach dem Fehler unterstützen können, und beenden Sie anschließend die Anwendung.

RC_SE_CONFLICT_WITH_INDEX_TYPE

Erläuterung: Die linguistische Angabe für den Suchbegriff der Abfrage entspricht nicht dem Indextyp. Beispielsweise kann PRECISE FORM OF nicht mit einem linguistischen Index verwendet werden. Der Standardwert für die linguistische Angabe wird verwendet, wie in Tabelle 7 auf Seite 198 gezeigt.

Aktion: Passen Sie Ihre Anwendung an, um Angaben von Abfrageoptionen zu vermeiden, die mit dem Indextyp in Konflikt stehen.

RC_SE_DICTIONARY_NOT_FOUND

Erläuterung: Die linguistischen Services des Text Extender können die Wörterverzeichnisdateien nicht finden. Die Abfrage wird ohne linguistische Unterstützung verarbeitet. Die entsprechenden Wörterverzeichnisdateien für die angegebenen Sprachencodes sind nicht im erwarteten Pfad.

Aktion: Sie können weiter API-Aufrufe durchführen. Für UNIX prüfen Sie, daß sich das erforderliche Wörterverzeichnis im Pfad `{DB2TX_INSTOWNERHOMEDIR}/db2tx/dicts` befindet. Für OS/2 prüfen Sie, daß sich das erforderliche Wörterverzeichnis in dem Pfad befindet, der in der Textkonfigurationseinstellung `DB2TX_DATA` angegeben ist. Falls erforderlich, installieren Sie das erforderliche Wörterverzeichnis.

RC_SE_DOCMOD_READ_PROBLEM

Erläuterung: Wenn ein Text Extender-Exemplar erstellt wird, wird eine Dokumentmodelldatei mit dem Namen `desmodel.ini` in das Exemplarverzeichnis gestellt. Wenn Sie einen Index erstellen, wird eine Datei `desmodel.ini` in das Indexverzeichnis `IXnnnnnn` gestellt. Diese Dokumentmodelldatei konnte nicht gelesen werden.

Aktion: Überprüfen Sie, daß eine Dokumentmodelldatei existiert und daß Sie im korrekten Verzeichnis steht.

RC_SE_DOCUMENT_NOT_ACCESSIBLE

Erläuterung: Das angeforderte Textdokument wurde gefunden, aber es kann momentan nicht darauf zugegriffen werden.

Aktion: Überprüfen Sie, ob das Dokument exklusiv im Zugriff einer anderen Task oder eines anderen Benutzers ist.

RC_SE_DOCUMENT_NOT_FOUND

Erläuterung: Das angeforderte Textdokument wurde nicht gefunden. Höchstwahrscheinlich wurde das Textdokument aus dem Speicher gelöscht, wurde aber noch nicht aus dem Text Extender-Index entfernt. Dieser Fehler kann auch auftreten, wenn Sie versuchen, ein Dokument anzuzeigen, das durch eine beschädigte Kennung identifiziert wird.

Aktion: In den meisten Fällen können Sie diesen Rückkehrcode ignorieren. Er wird nach der nächsten Indexaktualisierung nicht mehr angezeigt.

Bleibt das Problem bestehen, überprüfen Sie, daß Ihr Anwendungsprogramm die gefundene Kennung korrekt für das Anzeigen übergibt.

RC_SE_EMPTY_INDEX

Erläuterung: Der Text Extender-Index, der der Kennungsspalte zugeordnet ist, die durch die Suchanforderung adressiert wurde, ist leer. Entweder wurden keine Textdokumente zu diesem Index hinzugefügt oder alle Textdokumente wurden aus dem Index entfernt.

Dieser Fehler kann auftreten, wenn eine Textspalte aktiviert wurde, aber die Dokumente in der Spalte noch nicht indexiert wurden. Das heißt, Sie haben im Befehl ENABLE TEXT COLUMN angegeben, den Index später, zu einem Zeitpunkt, der durch die Einstellung für das regelmäßige Indexieren festgelegt ist, zu erstellen.

Dieser Fehler kann auch auftreten, wenn eine Texttabelle aktiviert wurde, einen leeren gemeinsamen Index für alle Textspalten zu erstellen, aber keine der Textspalten aktiviert wurde.

Aktion: Wenn ENABLE TEXT TABLE verwendet wurde, um einen leeren gemeinsamen Index für alle Textspalten zu erstellen, führen Sie ENABLE TEXT COLUMN mindestens für eine der Textspalten aus, die zu durchsuchenden Text enthält. In diesem Befehl können Sie festlegen, ob der Index sofort erstellt werden soll oder zu einem Zeitpunkt, der durch die Einstellungen für das regelmäßige Indexieren festgelegt ist.

Führen Sie GET INDEX STATUS aus, um zu überprüfen, daß der Index erfolgreich erstellt wurde.

RC_SE_EMPTY_QUERY

Erläuterung: Die angegebenen Suchkriterien wurden durch den Text Extender analysiert und linguistisch verarbeitet. Entweder führte ein Programmierfehler dazu, daß eine Abfrage durchgeführt wurde, die keine Suchbegriffe enthielt, oder alle Suchbegriffe waren Stoppwörter (Wörter, die vom Text Extender nicht indexiert werden), die aus der Abfrage entfernt werden. Als Ergebnis ergaben sich keine Suchbegriffe.

Aktion: Formulieren Sie die Abfrage um. Bleibt das Problem bestehen, suchen Sie nach einem Programmierfehler.

RC_SE_END_OF_INFORMATION

Erläuterung: Dies ist kein Fehler. Das Ende des Dokuments wurde erreicht. Für DesGetMatches stehen keine weiteren Informationen zur Verfügung.

Aktion: Verwenden Sie diesen Rückkehrcode, um die iterative Verarbeitung des Dokuments mit DesGetMatches zu beenden.

RC_SE_FUNCTION_DISABLED

Erläuterung: Die angeforderte Funktion hat eine Text Extender-Funktion aufgerufen, die durch den Administrator verhindert wurde.

Aktion: Bitten Sie Ihren Administrator um Unterstützung. Es ist möglicherweise erforderlich, den Text Extender zu stoppen und erneut zu starten (`txstop`/`txstart`).

RC_SE_FUNCTION_IN_ERROR

Erläuterung: Die angeforderte Funktion wurde gesperrt aufgrund einer Fehlerbedingung auf dem Text Extender-Server. Der API-Aufruf kann nicht verarbeitet werden.

Aktion: Prüfen Sie den Indexstatus. Prüfen Sie den verfügbaren Speicherbereich im Indexverzeichnis. Setzen Sie den Indexstatus zurück und wiederholen Sie den Befehl.

RC_SE_INCORRECT_HANDLE

Erläuterung: Eine Kennung, die in einem Eingabeparameter angegeben wurde, z.&rbl;B. *anzeigesitzungskennung*, ist nicht gültig. Es muß sich um eine Kennung handeln, die von einem vorherigen Aufruf zurückgegeben wurde und die nicht veraltet ist.

Aktion: Speichern Sie alle Informationen, die Sie bei der Suche nach der Fehlerursache unterstützen können, und beenden Sie anschließend die Anwendung durch Aufrufen von DesEndBrowseSession.

Überprüfen Sie, ob ein Programmierfehler eine falsche Kennung erstellt.

RC_SE_INDEX_DELETED

Erläuterung: Der Text Extender-Index, auf den zugegriffen werden soll, ist gelöscht.

Aktion: Kontaktieren Sie den Text Extender-Administrator, um den Index erneut zu erstellen.

RC_SE_INDEX_NOT_ACCESSIBLE

Erläuterung: Auf den Text Extender-Index kann nicht zugegriffen werden und der aktuelle Aufruf kann nicht verarbeitet werden.

Aktion: Bitten Sie den Text Extender-Administrator, die Zugriffsmöglichkeit auf den Index zu prüfen.

RC_SE_INDEX_SUSPENDED

Erläuterung: Der Text Extender hat eine Anforderung in bezug auf einen Text Extender-Index empfangen, der in einer anderen Sitzung oder in der aktuellen Sitzung ausgeschlossen wurde.

Aktion: Bitten Sie den Text Extender-Administrator, den Status des Indexes zu prüfen.

RC_SE_INSTALLATION_PROBLEM

Erläuterung: Der Text Extender ist auf einen Installationsfehler gestoßen.

Aktion: Überprüfen Sie die aktuelle Einstellung der Umgebungsvariablen DB2INSTANCE, DB2TX_INSTOWNER und DB2TXINSTOWNERHOMEDIR. Verwenden Sie `descfgcl -d` und `descfgsv -d -i txins`*`nnn`*, um Ihre Suchservicekonfiguration zu überprüfen.

RC_SE_IO_PROBLEM

Erläuterung: Als der Server versuchte, eine seiner Indexdateien zu öffnen oder zu lesen, ist ein Fehler aufgetreten. Folgende Gründe können dafür vorliegen:

Eine nicht beabsichtige Aktion durch den Administrator, z.&rbl;B. das Löschen einer Text Extender-Indexdatei

Falsche Einstellung in der Textkonfiguration

Aktion: Beenden Sie die Anwendung. Prüfen Sie mit dem Administrator, ob alle Dateien des aktuellen Text Extender-Indexes existieren und ob die Textkonfigurationseinstellungen korrekt sind.

RC_SE_LS_FUNCTION_FAILED

Erläuterung: Eine Funktion, die auf die Datenbank zugegriffen hat, um Textdokumente für die Anzeige abzurufen, ist fehlgeschlagen. Entweder kann der Benutzer nicht länger auf die Datenbank zugreifen oder der Benutzer ist für die Texttabelle nicht berechtigt.

Aktion: Überprüfen Sie, daß die Eingabe für die Funktion, z.&rbl;B. die Benutzer-ID, korrekt ist. Überprüfen Sie, daß auf die Datenbank zugegriffen werden kann und daß der Benutzer für die Task berechtigt ist.

RC_SE_LS_NOT_EXECUTABLE

Erläuterung: Eine Funktion, die versuchte, auf die Datenbank zuzugreifen, um Textdokumente für die Anzeige abzurufen, kann nicht ausgeführt werden.

Aktion: Überprüfen Sie, ob der Text Extender korrekt installiert ist. Bleibt das Problem bestehen, kontaktieren Sie Ihren IBM Ansprechpartner.

RC_SE_MAX_OUTPUT_SIZE_EXCEEDED

Erläuterung: Eine ungewöhnlich hohe Anzahl an Übereinstimmungen wurde gefunden. Die Größe der Anzeigeinformationen hat den Maximalwert, der verarbeitet werden kann, überschritten. Die Anforderung kann nicht verarbeitet werden.

Aktion: Grenzen Sie die Abfrage weiter ein oder stellen Sie sicher, daß mehr Systemspeicher verfügbar ist.

RC_SE_MAX_NUMBER_OF_BUSY_INDEXES

Erläuterung: Die angeforderte Funktion wurde durch den Suchservice verhindert, da die maximale Anzahl von Indizes momentan aktiv ist.

Aktion: Geben Sie nach einem kurzen Zeitraum den Funktionsaufruf erneut aus. Im allgemeinen herrscht das Problem nur temporär.

RC_SE_NO_DATA

Erläuterung: Dies ist kein Fehler. Kein Textdokument stimmt mit den Suchkriterien überein. Wenn Sie Anzeigeinformationen anfordern, werden keine Anzeigeinformationen zurückgegeben. Für die Anzeigeinformationen wird kein Speicher zugeordnet.

Aktion: Keine Aktion erforderlich.

RC_SE_NOT_ENOUGH_MEMORY

Erläuterung: Zu wenig Speicherplatz auf dem Client oder auf dem Server-System. Die aktuelle Anforderung kann nicht verarbeitet werden.

Aktion: Geben Sie Speicherbereich frei und beenden Sie die Anwendung.

RC_SE_PROCESSING_LIMIT_EXCEEDED

Erläuterung: Die aktuelle Suchanforderung hat die maximale Ergebnisgröße oder die maximale Verarbeitungszeit, die für Ihre Client-/Server-Umgebung angegeben ist, überschritten. Die Anforderung wurde abgebrochen.

Aktion: Grenzen Sie die Suchanforderung weiter ein. Erhöhen Sie gegebenenfalls die maximale Verarbeitungszeit.

RC_SE_QUERY_TOO_COMPLEX

Erläuterung: Die angegebene Abfrage ist zu komplex.

Aktion: Passen Sie Ihre Anwendung an, um zu verhindern, daß zu viele Platzhalterzeichen und Synonyme verwendet werden.

Die übermäßige Verwendung von Platzhalterzeichen oder der Option SYNONYM können eine Abfrage auf eine Größe erweitern, die vom Text Extender nicht verwaltet werden kann.

RC_SE_REQUEST_IN_PROGRESS

Erläuterung: Ein Text Extender-API-Anzeigeservice wurde aufgerufen, während eine andere API-Anzeigeanforderung für dieselbe Sitzung aktiv war.

Aktion: Beenden Sie die Sitzung durch Aufrufen von DesEndBrowseSession und geben Sie Speicherplatz frei, indem Sie DesFreeBrowseInfo aufrufen.

Der Text Extender-API-Anzeigeservice unterstützt den gleichzeitigen Zugriff auf dieselbe Anzeigesitzung nicht.

Alle Anwendungen, die gleichzeitig in demselben Prozeß aktiv sind, sollten ihre eigene Anzeigesitzung bearbeiten.

RC_SE_SERVER_BUSY

Erläuterung: Der Text Extender-Client kann derzeit keine Sitzung mit dem angeforderten Text Extender-Server aufbauen, oder die Text Extender-Server-Kommunikationsverbindung wurde unterbrochen und kann nicht erneut hergestellt werden.

Der Text Extender-Server wurde korrekt gestartet, aber die maximale Anzahl der parallelen Server-Prozesse wurde erreicht.

Aktion: Wenn dies kein temporäres Problem ist, ändern Sie die Kommunikationskonfiguration auf dem Text Extender-Server.

RC_SE_SERVER_CONNECTION_LOST

Erläuterung: Die Verbindung zwischen Client und Server wurde unterbrochen und kann nicht erneut hergestellt werden.

Die Text Extender-Server-Task wurde möglicherweise durch einen Administrator gestoppt, oder die Server-Workstation wurde möglicherweise heruntergefahren.

Aktion: Prüfen Sie, ob eine dieser Bedingungen aufgetreten ist, und lassen Sie sie korrigieren.

RC_SE_SERVER_NOT_AVAILABLE

Erläuterung: Die Text Extender-API-Services konnten keine Sitzung mit dem angeforderten Text Extender-Server aufbauen.

Der Text Extender-Server wurde möglicherweise nicht gestartet.

Aktion: Überprüfen Sie, daß der Text Extender-Server korrekt gestartet wurde. Bleibt der Fehler bestehen, handelt es sich um einen Installationsfehler.

RC_SE_STOPWORD_IGNORED

Erläuterung: Dieser Informationscode wird zurückgegeben, wenn die angegebene Abfrage mindestens einen Suchbegriff enthält, der nur aus Stoppwörtern besteht. Der Suchbegriff wurde bei der Verarbeitung der Abfrage ignoriert.

Aktion: Sie können weiter API-Aufrufe ausgeben. Vermeiden Sie die Verwendung von Stoppwörtern in Text Extender-Abfragen.

RC_SE_UNEXPECTED_ERROR

Erläuterung: Ein Fehler ist aufgetreten, der möglicherweise durch eine inkorrekte Installation des Text Extender verursacht wurde.

Aktion: Beenden Sie die Anwendung, wobei Sie alle Informationen, die Sie bei der Suche nach der Fehlerursache unterstützen können, speichern.

RC_SE_UNKNOWN_INDEX_NAME

Erläuterung: Der Name des Textindexes, der einer Textspalte zugeordnet ist, ist Teil einer Kennung.

Aktion: Stellen Sie sicher, daß die Kennung, die Sie als Eingabe für DesGetBrowseInfo verwenden, korrekt ist.

RC_SE_UNKNOWN_SECTION_NAME

Erläuterung: Ein angegebener Abschnittsname ist nicht Teil eines Modells, das in einer Dokumentmodelldatei angegeben ist, bzw. nicht Teil des verwendeten Standardmodells.

Aktion: Geben Sie einen Abschnittsnamen an, der Teil des angegebenen Modells oder des Standardmodells ist.

RC_SE_UNKNOWN_SERVER_NAME

Erläuterung: Der Name des Text Extender-Servers ist Teil einer Kennung.

Aktion: Stellen Sie sicher, daß die Kennung, die Sie als Eingabe für DesGetBrowseInfo verwenden, korrekt ist.

RC_SE_WRITE_TO_DISK_ERROR

Erläuterung: Ein Schreibfehler ist aufgetreten, der möglicherweise durch eine volle Platte auf der Text Extender-Server-Workstation oder durch eine inkorrekte Installation des Text Extender verursacht wurde.

Aktion: Beenden Sie die Anwendung, wobei Sie alle Informationen, die Sie bei der Suche nach der Fehlerursache unterstützen können, speichern. Überprüfen Sie, daß auf dem Server ausreichend Plattenspeicherplatz verfügbar ist.

RC_SQL_ERROR_WITH_INFO

Erläuterung: SQL-Fehler aufgetreten. Eine Fehlernachricht wird zurückgegeben.

Aktion: Prüfen Sie die Fehlernachricht, die vom Text Extender zurückgegeben wurde, nach weiteren Informationen, z.&rbl;B. SQL-Fehlernachricht, SQLState und Basis-SQL-Fehlercode.

RC_SQL_ERROR_NO_INFO

Erläuterung: SQL-Fehler aufgetreten. Es wird keine Fehlernachricht zurückgegeben.

RC_TEXT_COLUMN_NOT_ENABLED

Erläuterung: Die angegebene Kennungsspalte ist keine Spalte in der von Ihnen angegebenen Tabelle.

Aktion: Überprüfen Sie, daß der Kennungsspaltenname, den Sie angegeben haben, korrekt ist. Stellen Sie sicher, daß die Textspalte in dieser Tabelle aktiviert wurde.

Kapitel 13. Nachrichten

In diesem Kapitel werden folgende Themen beschrieben:

- **SQLSTATES, die von UDFs zurückgegeben werden:** Diese Nachrichten können angezeigt werden, wenn Sie UDFs verwendet.
- **Nachrichten vom DB2TX-Befehlszeilenprozessor:** Diese Nachrichten können angezeigt werden, wenn Sie unter Verwendung des Befehlszeilenprozessors DB2TX Verwaltungsbefehle eingeben. Vor jeder Nachrichtennummer steht das Präfix DES.

SQLSTATES, die von UDFs zurückgegeben werden

Die benutzerdefinierten Funktionen, die vom Text Extender geliefert werden, können Fehlernachrichten zurückgeben. Beispiel:

```
SQL0443N User-defined function
"DB2TX.CONTAINS" (specific name "DES5A")
has returned an error SQLSTATE with
diagnostic text "Nachrichtendatei kann nicht geöffnet werden".
SQLSTATE=38702
```

Die Nachrichten in diesem Abschnitt sind nach SQLSTATE-Nummer sortiert.

01H10 Die Datei *dateiname* kann nicht geöffnet werden.

Aktion: Stellen Sie sicher, daß die Datei existiert und daß der DB2-Exemplarname über die notwendigen Berechtigungen verfügt, sie zu öffnen.

01H11 Eine Textkennung ist unvollständig.

Erläuterung: Ein Versuch wurde unternommen, eine Kennung zu verwenden, die initialisiert, aber nicht abgeschlossen wurde. Eine Teilkennung wurde unter Verwendung von INIT_TEXT_HANDLE erstellt, die vordefinierte Werte für Sprache und Format des Dokuments enthält. Die Kennung wurde jedoch nicht durch einen Auslöser abgeschlossen.

Aktion: Verwenden Sie nur Kennungen, die abgeschlossen sind. Wenn die betreffende Kennung in einer Kennungsspalte gespeichert ist, aktivieren Sie deren entsprechende Textspalte (erneut).

01H12 Suchargumente sind zu lang. Das zweite Argument wurde ignoriert.

Erläuterung: Die UDF REFINE wurde verwendet, um zwei Suchargumente zu kombinieren, aber die kombinierte Länge der Suchargumente ist größer als die maximal für LONG VARCHAR zulässige Länge. Die UDF REFINE gibt das erste Suchargument anstelle des kombinierten Suchargumentes zurück.

Aktion: Verringern Sie die Länge eines oder beider Suchargumente und wiederholen Sie anschließend die Abfrage.

01H13 Suchargument enthält ein Stoppwort, bzw. Suchargumente enthalten Stoppwörter.

Erläuterung: Die angegebene Abfrage enthält mindestens einen Suchbegriff, der nur aus Stoppwörtern besteht. Der Suchbegriff wurde bei der Verarbeitung der Abfrage ignoriert.

Aktion: Vermeiden Sie die Verwendung von Stoppwörtern in Text Extender-Abfragen.

01H14 Ein Wörterverzeichnis einer Sprache für die Verarbeitung auf linguistischer Basis fehlt.

Erläuterung: Die linguistischen Services des Text Extender können die Wörterverzeichnisdateien nicht finden. Die Abfrage wird ohne linguistische Unterstützung verarbeitet. Die entsprechenden Wörterverzeichnisdateien für die angegebenen Sprachencodes sind nicht im erwarteten Pfad.

Aktion: Für UNIX prüfen Sie, daß sich das erforderliche Wörterverzeichnis im Pfad `{DB2TX_INSTOWNERHOMEDIR}/db2tx/dicts` befindet. Für OS/2 prüfen Sie, daß sich das erforderliche Wörterverzeichnis in dem Pfad befindet, der in den Textkonfigurationseinstellungen angegeben ist. Falls erforderlich, installieren Sie das erforderliche Wörterverzeichnis.

01H15 Die Angabe eines linguistischen Suchbegriffs entspricht nicht dem Indextyp.

Erläuterung: Die linguistische Angabe für den Suchbegriff der Abfrage entspricht nicht dem Indextyp. Beispielsweise sollte PRECISE FORM OF nicht mit einem linguistischen Index verwendet werden. Der Standardwert für die linguistische Angabe wird verwendet, wie in Tabelle 7 auf Seite 198 gezeigt.

Aktion: Passen Sie Ihre Anwendung an, um Angaben von Abfrageoptionen zu vermeiden, die mit dem Indextyp in Konflikt stehen.

38700 Die Text Extender-Bibliothek ist nicht aktuell.

Erläuterung: Ein Versuch wurde unternommen, eine Kennung zu verwenden, die nur von einer späteren Version der Text Extender-Bibliothek interpretiert werden kann.

Aktion: Stellen Sie sicher, daß der Pfad zur aktuellen Bibliotheksversion korrekt gesetzt ist und daß Sie über die notwendigen Berechtigungen verfügen, darauf zuzugreifen.

Suchen Sie in der DB2-Katalogsicht SYSCAT.FUNCTIONS in der Spalte IMPLEMENTATION nach der UDF, die den Fehler verursacht hat.

38701 *trace_datei* **Diese Trace-Datei kann nicht geöffnet werden.**

Erläuterung: Ein Versuch wurde unternommen, eine Trace-Funktion zu verwenden, die in die Datei DB2TX_TRACEFILE im Verzeichnis DB2TX_TRACEDIR schreibt. Entweder existiert die Datei nicht, kann sie nicht gefunden werden, oder die notwendigen Berechtigungen für die Datei sind nicht verfügbar.

38702 Nachrichtendatei *nachrichtendatei* **kann nicht geöffnet werden.**

Erläuterung: Eine Situation ist aufgetreten, die dazu führte, daß der Text Extender versucht hat, eine Nachricht zurückzugeben. Entweder existiert die Datei, die die Nachrichten enthält, nicht, kann sie nicht gefunden werden, oder die notwendigen Berechtigungen für die Datei sind nicht verfügbar.

Aktion: Stellen Sie sicher, daß die Datei existiert, daß der Pfad korrekt gesetzt ist und daß Sie über die notwendigen Berechtigungen verfügen, die Datei zu öffnen.

38704 Das Format der Textkennung ist falsch.

Erläuterung: Eine Kennung mit einem falschen Format wurde als Argument für eine Text Extender-UDF verwendet.

Aktion: Stellen Sie sicher, daß die Kennung nicht durch INIT_TEXT_HANDLE erstellt wurde.

38705 *udf_name* **Falsche Deklaration der benutzerdefinierten Funktion.**

Erläuterung: Der spezifische Name einer UDF wurde in der Prozedur, in der UDFs deklariert werden, geändert. UDF-Namen können geändert werden, aber nicht deren spezifische Namen.

Aktion: Überprüfen Sie die Prozedur DESCVDF.DDL, die die UDF-Deklarationen enthält, um sicherzustellen, daß noch immer die korrekten Namen verwendet werden. Vergleichen Sie die Namen mit denen im Originalprogrammdatenträger.

38706 *attribut* **Dieses Attribut kann nicht erkannt werden.**

Erläuterung: Ein Versuch wurde unternommen, eine CCSID, ein Format oder eine Sprache auf einen unbekannten Wert zu setzen.

Aktion: Die korrekten Werte finden Sie in „Informationen zu Textdokumenten“ auf Seite 256.

38708 *rückkehrcode* **Suche ist fehlgeschlagen.**

Erläuterung: Ein Fehler ist während der Verarbeitung einer Suchanforderung aufgetreten.

Aktion: Eine Beschreibung des Rückkehrcodes finden Sie in Kapitel 12, „Rückkehrcodes“ auf Seite 225.

38709 Nicht genügend Speicher verfügbar.

Erläuterung: Zum Ausführen der UDF steht nicht genügend Speicher zur Verfügung.

Aktion: Schließen Sie unnötige Anwendungen, um Speicher freizugeben, und wiederholen Sie den Versuch.

38710 *fehlernummer* **Kein Zugriff auf Suchergebnisse.**

Erläuterung: Ein Fehler ist aufgetreten, während versucht wurde, die Liste der gefundenen Dokumente (Ergebnisliste) zu lesen, die vom Suchservice zurückgegeben wurde.

Aktion: Wiederholen Sie die Suche. Ist dies nicht erfolgreich, starten Sie den Suchservice erneut. Bleibt das Problem bestehen, melden Sie es an Ihren lokalen IBM Ansprechpartner und geben Sie dabei die Fehlernummer an.

38711 Schwerwiegender interner Fehler.

Erläuterung: Ein schwerwiegender Fehler ist aufgetreten.

Aktion: Melden Sie den Fehler an Ihren lokalen IBM Ansprechpartner und geben Sie dabei die Umstände an, unter denen er aufgetreten ist.

38712 *indexname* **Falsche Textkennung in diesem Textindex.**

Erläuterung: Eine Kennung wurde beschädigt.

Aktion: Verwenden Sie UPDATE INDEX, um den Index wiederherzustellen.

38714 Die Umgebungsvariable DB2TX_INSTOWNERHOMEDIR kürzen.

Erläuterung: Der Name des Ausgangsverzeichnisses des Exemplareigners darf nicht länger als 256 Zeichen sein.

Aktion: Verwenden Sie Verbindungen, um die Länge des Verzeichnisnamens zu verringern.

Nachrichten vom Text Extender

Jede Nachricht hat eine Nachrichten-ID, die aus einem Präfix (DES), der Nachrichtennummer und einem Suffix besteht. Das Suffix (ein Buchstabe) gibt an, wie schwerwiegend das Ereignis war, das die Nachricht hervorgerufen hat:

I Informationsnachricht

W Warnung

N Fehlernachricht (oder “negative” Nachricht)

C Nachricht zu einem kritischen Fehler

DES0001N Ungültige Anzahl an Argumenten.

Erläuterung: Der Befehl `db2txinstance` benötigt zwei Argumente.

Aktion: Geben Sie den Befehl erneut mit diesen Argumenten an:

```
db2txinstance exemplarname db2_exemplarname
```

Dabei ist *exemplarname* der Anmeldename eines bestehenden UNIX-Benutzers, der als Eigner dieses Exemplar zugeordnet ist, und *db2_exemplarname* ist der Anmeldename des Eigners des entsprechenden DB2-Exemplars.

DES0002N Ungültiger Exemplarname.

Erläuterung: Der angegebene Exemplarname muß der Anmeldename eines bestehenden UNIX-Benutzers sein.

Aktion: Korrigieren Sie den Exemplarnamen, wählen Sie einen bestehenden UNIX-Benutzer aus oder erstellen Sie einen UNIX-Benutzer, der Exemplareigner sein soll.

Geben Sie den Befehl `db2txinstance` erneut wie folgt ein:

```
db2txinstance exemplarname
```

Dabei ist *exemplarname* der Anmeldename des ausgewählten UNIX-Benutzers.

DES0004N Das angegebene Exemplar existiert bereits. Der Befehl kann nicht verarbeitet werden.

Erläuterung: *exemplareigner* gibt den Anmeldenamen eines UNIX-Benutzers an, der der Eigner des Exemplars ist. Dieser Exemplareigner hat bereits ein Verzeichnis `db2tx` im Ausgangsverzeichnis.

Aktion: Um das Exemplar zu erstellen, löschen Sie das bestehende Exemplar und wiederholen Sie den Befehl.

DES0005N Der Katalog der Installationsnachrichten kann nicht gefunden werden.

Erläuterung: Der Nachrichtenkatalog, der für die Installationsprozeduren erforderlich ist, fehlt auf dem System; möglicherweise wurde er gelöscht, oder die Datenbankprodukte wurden möglicherweise falsch geladen.

Aktion: Prüfen Sie, daß die Produktkomponente db2tx_01_01_0000.client korrekt installiert ist. Wenn Fehler bei der Prüfung auftreten, die Komponente erneut installieren.

DES0015W Die Angabe eines linguistischen Suchbegriffs entspricht nicht dem Indextyp.

Erläuterung: Die linguistische Angabe für den Suchbegriff der Abfrage entspricht nicht dem Indextyp. Beispielsweise sollte PRECISE FORM OF nicht mit einem linguistischen Index verwendet werden. Der Standardwert für die linguistische Angabe wird verwendet, wie in Tabelle 7 auf Seite 198 gezeigt.

Aktion: Passen Sie Ihre Anwendung an, um Angaben von Abfrageoptionen zu vermeiden, die mit dem Indextyp in Konflikt stehen.

DES0016W Eine Sprachangabe wurde für den Indextyp nicht unterstützt.

Erläuterung: Die von Ihnen angegebene Sprache wird für den angegebenen Indextyp nicht unterstützt.

Aktion: Eine Liste der unterstützten Sprachen für den Indextyp finden Sie in der Dokumentation.

DES0017W Die Merkmalextraktion wurde nicht aktiviert.

Erläuterung: Sie haben ein Merkmalsuchargument in Ihrer Abfrage verwendet, aber der Index wurde nicht mit der Indexoption FEATURE_EXTRACTION erstellt.

Aktion: Ändern Sie die Indexoption in FEATURE_EXTRACTION.

DES0018W *option* **wird für den aktuellen Indextyp nicht unterstützt.**

Erläuterung: Sie haben eine Suchoption angefordert, die für den aktuellen Indextyp und die Indexoption nicht unterstützt wird.

Aktion: Prüfen Sie, welcher Indextyp oder welche Indexoption die angeforderte Suchoption unterstützt. Siehe Tabelle 7 auf Seite 198.

DES0121N Es konnte kein Speicher zugeordnet werden.

Erläuterung: Für die Anwendung konnte kein Speicher reserviert werden.

Aktion: Erhöhen Sie den Seitenwechselbereich.

DES0333N Im Textindex ist ein Datei-E/A-Fehler aufgetreten.

Erläuterung: Der Text Extender-Client kann keine Sitzung mit dem angeforderten Server aufbauen.

Aktion: Überprüfen Sie, daß der Text Extender-Server gestartet wurde. Andernfalls führen Sie TXSTART aus.

DES0709W Das Wörterverzeichnis für die Sprache *sprache* **ist nicht installiert.**

Erläuterung: Der Text Extender kann die Wörterverzeichnisdateien nicht finden.

Aktion: Installieren Sie das Wörterverzeichnis für die angegebene Sprache (erneut).

DES0377N Im Textindex ist ein Datei-E/A-Fehler aufgetreten.

Erläuterung: Der Text Extender kann auf den Textindex nicht zugreifen. Dies kann passieren, wenn die Einstellung DIRECTORY in der Textkonfiguration auf ein ungültiges Verzeichnis zeigt.

Aktion: Überprüfen Sie die Textkonfigurationseinstellungen.

DES0700N Der Zahlenwert *zahl* **für den Knoten ist nicht in der Knotengruppendefinition enthalten.**

Erläuterung: Die angegebene Knotennummer ist ungültig.

Aktion: Überprüfen Sie die DB2-Knotennummer.

DES0701N Der Zahlenwert *zahl* **für den Knoten ist außerhalb des gültigen Bereichs.**

Erläuterung: Die angegebene Knotennummer ist ungültig.

Aktion: Überprüfen Sie die DB2-Knotennummer.

DES0704N Für das Format *format* **ist die Angabe eines Indexmerkmals erforderlich.**

Erläuterung: Das Dokumentformat ist nicht mit den Informationen zum Indextyp kompatibel.

Aktion: Geben Sie ein Indexmerkmal an, das mit dem Dokumentformat kompatibel ist.

DES0705N Der angegebene Dokumentmodellname *name* **wurde in der Modelldefinitionsdatei nicht gefunden.**

Erläuterung: Der Dokumentmodellname wurde in der Modelldefinitionsdatei nicht gefunden. Berücksichtigen Sie, daß bei dem Modellnamen die Groß-/Kleinschreibung beachtet wird.

Aktion: Verwenden Sie einen Modellnamen, der in der Modelldefinitionsdatei angegeben ist.

DES0706N Auf die Modelldefinitionsdatei auf dem DB2 Text Extender-Server kann nicht zugegriffen werden.

Erläuterung: Die Modelldefinitionsdatei konnte nicht gefunden oder nicht geöffnet werden.

Aktion: Überprüfen Sie, daß die Modelldefinitionsdatei existiert.

DES0707N Das Format *format* **unterstützt das angegebene Indexmerkmal nicht.**

Erläuterung: Das Dokumentformat unterstützt das Indexmerkmal nicht.

Aktion: Geben Sie ein Indexmerkmal an, das mit dem Dokumentformat kompatibel ist.

DES0710N Ein Nullzeiger ist für den Parameter *parameter* **nicht zulässig.**

Erläuterung: Für den Parameter *parameter* ist kein Wert angegeben.

Aktion: Geben Sie einen Wert für den Parameter *parameter* an.

DES0711N Ein interner Fehler im Text Extender ist aufgetreten. Diagnoseinformationen: *infos***.**

Erläuterung: Ein interner Verarbeitungsfehler ist aufgetreten.

Aktion: Prüfen Sie die Diagnosenachricht, um das Problem zu lösen. Wenn der interne Fehler nicht durch einen Installationsfehler verursacht wurde, finden Sie möglicherweise in der Datei desdiag.log oder in einer erstellten Trace-Datei zusätzliche Informationen. Hilft dies nicht, sammeln Sie die verfügbaren Informationen und rufen Sie den IBM Kundendienst an.

DES0712N Parameter *parameter* **ist zu lang.**

Erläuterung: Der angegebene Parameter *parameter* ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie den Parameter *parameter* unter Verwendung einer gültigen Länge an.

DES0713N Die Länge des Parameters *parameter* **ist falsch:** *länge***.**

Erläuterung: Der angegebene Parameter *parameter* ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie den Parameter *parameter* unter Verwendung einer gültigen Länge an.

DES0714N Parameter *parameter* **entweder direkt oder über eine Umgebungsvariable angeben.**

Erläuterung: Die CCSID, das Format oder die Sprache wurde nicht angegeben, und für diesen Wert ist keine Textkonfigurationseinstellung vorhanden.

Aktion: Geben Sie den fehlenden Parameter entweder direkt im Befehl ENABLE TEXT COLUMN an oder setzen Sie einen Wert in den Textkonfigurationseinstellungen.

DES0715N Datentyp *schema.typ* **wird für Textdaten nicht unterstützt.**

Erläuterung: *schema.typ* ist der Schemaname und der Typenname der Textspalte oder das Ergebnis einer Zugriffsfunktion. Der Datentyp für eine Textspalte wird vom Text Extender nicht unterstützt. Der Datentyp muß CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB sein. Ist dies nicht der Fall, müssen Sie eine Zugriffsfunktion liefern, deren Eingabewert der Datentyp der Textspalte und deren Ausgabewert VARCHAR, LONG VARCHAR oder LOB ist.

Aktion: Wenn *schema.typ* ein Textspaltentyp ist, müssen Sie eine Zugriffsfunktion mit einem Ergebnis vom Typ VARCHAR, LONG VARCHAR oder LOB registrieren. Wenn *schema.typ* das Ergebnis einer Zugriffsfunktion ist, kann es nicht verwendet werden. Stellen Sie eine Zugriffsfunktion mit einem Ergebnis vom erforderlichen Typ zu Verfügung.

DES0716N Format *format* **wird nicht unterstützt.**

Erläuterung: *format* ist ein Format, das vom Text Extender nicht unterstützt wird.

Aktion: Überprüfen Sie die Liste der unterstützten Formate in „Formate“ auf Seite 256.

DES0717N Sprache *sprache* **wird nicht unterstützt.**

Erläuterung: *sprache* ist eine Sprache, die vom Text Extender nicht unterstützt wird.

Aktion: Überprüfen Sie die Liste der unterstützten Sprachen in „Sprachen“ auf Seite 258.

DES0718N CCSID *ccsid_wert* **wird nicht unterstützt.**

Erläuterung: Sie haben einen ungültigen CCSID-Wert angegeben.

Aktion: Eine Liste der unterstützten CCSIDs finden Sie in der Dokumentation.

DES0719N Ein Aufruf des DB2 Text Extender-Programms *programm* **ist fehlgeschlagen. Rückkehrcode:** *rc***.**

Erläuterung: Während der Installation ist möglicherweise ein Fehler aufgetreten. Die Rückkehrcodes werden in der Datei DES_EXT.H aufgelistet.

Aktion: Überprüfen Sie, ob die Installation erfolgreich war. Überprüfen Sie, daß die Umgebungsvariablen, wie z.&rbl;B. DB2TX_INSTOWNER und DB2TX_INSTOWNERHOMEDIR, korrekt gesetzt sind.

DES0720N Die Zugriffsfunktion *schema.funktion* **ist in der Datenbank nicht registriert.**

Erläuterung: Der Name der Funktion ist entweder falsch oder er wurde nicht für die Datenbank registriert.

Aktion: Überprüfen Sie den Namen der Zugriffsfunktion. Wenn Sie korrekt ist, überprüfen Sie, daß die Funktion dem Datenbanksystem bekannt ist. Verwenden Sie CREATE FUNCTION, um die Zugriffsfunktion für die Datenbank zu registrieren.

DES0721N Die Datenbank ist inkonsistent; eine DB2 Text Extender-Katalogsicht fehlt.

Erläuterung: Eine der Text Extender-Katalogsichten ist nicht in der Datenbank.

Aktion: Verwenden Sie den Befehl DISABLE DATABASE, um die verbleibenden Katalogsichten zu entfernen, und geben Sie anschließend ENABLE DATABASE erneut ein. Die Indexdaten sind verloren; indexieren Sie die Textdokumente erneut.

DES0722N Tabelle *schema.tabelle* **ist keine Basistabelle in der Datenbank.**

Erläuterung: Entweder existiert die Tabelle nicht in der Datenbank oder es ist eine Ergebnistabelle oder eine Sicht. Eine Textspalte muß in einer Basistabelle sein, bevor sie für den Text Extender aktiviert werden kann.

Aktion: Stellen Sie sicher, daß der Tabellenname korrekt ist und daß die Tabelle eine Basistabelle ist.

DES0723N Das (erneute) Erstellen eines Textindexes für die Kennungsspalte *kennungsspaltenname* **in Tabelle** *schema.tabelle* **ist fehlgeschlagen.**

Erläuterung: Für die Kennungsspalte konnte kein Textindex erstellt werden.

Aktion: Verwenden Sie `txstatus`, um den Status des Servers zu prüfen. Wenn die Services auf dem Server korrekt laufen, verwenden Sie DISABLE TEXT COLUMN oder DISABLE TEXT TABLE, um erneut einen konsistenten Status zu erhalten. Aktivieren Sie anschließend die Textspalte erneut, indem Sie ENABLE TEXT COLUMN oder ENABLE TEXT TABLE verwenden.

DES0724N Ein Eintrag in der Katalogsicht TextIndices für die Kennungsspalte *kennungspalte* **in Tabelle** *schema.tabelle* **fehlt.**

Erläuterung: Die Katalogsicht TextIndices ist beschädigt.

Aktion: Verwenden Sie DISABLE TEXT COLUMN oder DISABLE TEXT TABLE, um erneut einen konsistenten Status zu erhalten. Aktivieren Sie anschließend die Textspalte, indem Sie ENABLE TEXT COLUMN oder ENABLE TEXT TABLE verwenden.

DES0727N Spalte *spalte* **in Tabelle** *schema.tabelle* **ist bereits mit der aktuellen Angabe für die Indexfunktion aktiviert.**

Erläuterung: Diese Nachricht kann auftreten, wenn die Tabelle gelöscht und anschließend unter Verwendung derselben Textspalte erneut erstellt wurde, ohne zunächst die Spalte zu inaktivieren.

Aktion: Inaktivieren Sie die Spalte und wiederholen Sie den Versuch.

DES0728N Spalte *spalte* **in Tabelle** *schema.tabelle* **nicht enthalten.**

Erläuterung: Sie versuchen eine Textspalte zu aktivieren, die nicht existiert.

Aktion: Ändern Sie den Tabellennamen oder den Spaltennamen und wiederholen Sie den Versuch.

DES0729N Kennungsspalte *kennungsspalte* **ist in Tabelle** *schema.tabelle* **nicht enthalten.**

Erläuterung: Sie versuchen eine Kennungsspalte zu verwenden, die nicht existiert.

Aktion: Verwenden Sie den Befehl GET STATUS, um zu prüfen, ob die Kennungsspalte existiert und daß ihr Name korrekt angegeben wurde.

DES0730N Tabelle *schema.tabelle* **ist bereits als Tabelle mit gemeinsamem Index aktiviert.**

Erläuterung: Sie versuchen, eine Tabelle zu aktivieren, die bereits als Tabelle mit einem gemeinsam benutzten Index aktiviert wurde.

Aktion: Entweder fahren Sie fort, ohne die Tabelle zu aktivieren, oder Sie führen den Befehl DISABLE TEXT TABLE aus, um die Tabelle zu inaktivieren, bevor Sie sie erneut aktivieren.

DES0731N Tabelle *schema.tabelle* **ist für den DB2 Text Extender nicht aktiviert; sie kann nicht inaktiviert werden.**

Erläuterung: Sie versuchen, eine Tabelle zu inaktivieren, die nicht aktiviert worden ist.

Aktion: Überprüfen Sie den Tabellennamen.

DES0732N Wert nahe Position *position* **der Aktualisierungsfrequenz ist nicht korrekt;** *parameter* **wurde erwartet.**

Erläuterung: Die Angabe *parameter* für die Aktualisierungsfrequenz war nicht korrekt.

Aktion: Überprüfen Sie den Parameter für die Aktualisierungsfrequenz und geben Sie den Befehl erneut ein.

DES0733N Tabelle *schema.tabelle* **enthält eine aktivierte Spalte; sie kann nicht als Tabelle mit gemeinsamem Index aktiviert werden.**

Erläuterung: Diese Tabelle enthält eine Textspalte, die bereits ihren eigenen Index hat. Sie können nicht einen gemeinsamen Index für alle Textspalten erstellen, solange dieser individuelle Index existiert.

Aktion: Verwenden Sie DISABLE TEXT COLUMN, um die aktivierten Spalten zu inaktivieren, und geben Sie anschließend den Befehl ENABLE TEXT TABLE erneut ein.

DES0734N Die Kennungsspalte *kennungsspalte* **gehört zu einer Tabelle, die teilweise Textdaten enthält,** *schema.tabelle***; sie kann nicht separat inaktiviert werden.**

Erläuterung: Sie können nicht eine einzelne Textspalte in einer Tabelle inaktivieren, die als Tabelle, die teilweise Textdaten enthält, aktiviert wurde.

Aktion: Inaktivieren Sie die gesamte Tabelle, die teilweise Textdaten enthält.

kennungsspalte **ist bereits eine Kennungsspalte in Tabelle** *schema.tabelle***.**

Erläuterung: Sie versuchen, einen vorhandenen Kennungsspaltennamen zu verwenden.

Aktion: Geben Sie den Befehl erneut ein, indem Sie einen anderen Namen für die Kennungsspalte verwenden.

DES0737N Tabelle *schema.tabellenname* **ist als eine Tabelle mit gemeinsamem Index mit der STORAGE-Option** *storage_option* **aktiviert.**

Erläuterung: Es ist nicht möglich, eine Tabelle mit einem gemeinsam benutzten Index für externe Dateien zu aktivieren.

Aktion: Wenn Sie eine Tabelle für externe Dateien aktivieren wollen, verwenden Sie eine Mehrindextabelle.

DES0738N Die Zugriffsfunktion *schema.funktion* **hat falsche Parameter.**

Erläuterung: Die Eingabe- oder Ausgabeparameter von *schema.funktion* sind falsch.

- Nur ein Eingabeparameter ist zulässig und er muß vom Datentyp der zu aktivierenden Textspalte sein.
- Der Ausgabeparameter muß vom Typ CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB sein.

DES0739W Das Programm zur Indexaktualisierung für Tabelle *schema.tabelle***, Kennungsspalte** *kennungsspaltenname***, konnte nicht gestartet werden.**

Erläuterung: Das Programm, das Indizes aktualisiert, konnte nicht gestartet werden. Während der Installation ist möglicherweise ein Fehler aufgetreten.

Aktion: Überprüfen Sie, ob die Installation erfolgreich war. Überprüfen Sie, daß die Umgebungsvariablen, wie z.&rbl;B. DB2TX_INSTOWNER und DB2TX_INSTOWNERHOMEDIR, korrekt gesetzt sind.

DES0740N Kennungsspalte *kennungsspalte* **kann nur für eine Textspalte des Typs** *typ* **wiederverwendet werden.**

Erläuterung: Die Kennungsspalte wurde bereits für einen anderen Kennungsspaltentyp verwendet.

Aktion: Geben Sie einen neuen Kennungsspaltennamen an.

DES0741N Programm oder Datei *parameter* **wurde nicht gefunden oder konnte nicht gestartet werden.**

Erläuterung: Der Befehl ENABLE DATABASE oder DISABLE DATABASE konnte die Datei *parameter* nicht öffnen. Während der Installation ist möglicherweise ein Fehler aufgetreten.

Aktion: Überprüfen Sie, ob die Installation erfolgreich war.

DES0742N Diese Tabelle verwendet Spaltenindizes. Im Befehl eine Kennungsspalte angeben.

Erläuterung: Die angegebene Tabelle ist als Mehrindextabelle aktiviert. Um mit einer bestimmten Spalte zu arbeiten, geben Sie die zugehörige Kennungsspalte an.

Aktion: Geben Sie eine Kennungsspalte an.

DES0745N Der DB2TX-Exemplareigner *exemplareigner* **ist keine gültige Benutzer-ID.**

Erläuterung: Die Umgebungsvariable DB2TX_INSTOWNER enthält keine gültige Benutzer-ID.

Aktion: Korrigieren Sie die Umgebungsvariable.

DES0747N Die aktuelle CCSID wird für den Indextyp *indextyp* **nicht unterstützt.**

Erläuterung: Sie haben eine CCSID angegeben, die für den angeforderten Indextyp nicht unterstützt wird.

Aktion: Eine Liste der unterstützten CCSIDs finden Sie in der Dokumentation.

DES0748N Der Wert für die COMMIT-Zählung *wert* **wird vom DB2 Text Extender nicht unterstützt.**

Erläuterung: Der angegebene Wert für die COMMIT-Zählung wird nicht unterstützt.

Aktion: Geben Sie einen gültigen Wert für die COMMIT-Zählung an.

DES0749N Der Wert für den Aktualisierungsindex *wert* **ist beim DB2 Text Extender nicht bekannt.**

Erläuterung: Der angegebene Wert für den Aktualisierungsindex ist ungültig.

Aktion: Geben Sie einen gültigen Wert für den Aktualisierungsindex an.

DES0750N Der Wert für den Indextyp *indextyp* **ist beim DB2 Text Extender nicht bekannt.**

Erläuterung: Der angegebene Wert für den Indextyp ist ungültig.

Aktion: Geben Sie einen gültigen Wert für den Indextyp an.

DES0751N Keine Berechtigung zur Ausführung der angegebenen Operation.

Erläuterung: Sie verfügen nicht über die erforderliche Datenbankadministratorberechtigung, um diese Operation auszuführen.

Aktion: Lassen Sie diese Operation von einem Datenbankadministrator ausführen.

DES0756N Die Datenbank ist für DB2 Text Extender nicht aktiviert.

Erläuterung: Die Datenbank muß aktiviert sein, bevor dieser Befehl ausgeführt werden kann.

Aktion: Führen Sie ENABLE DATABASE aus und übergeben Sie den Befehl anschließend erneut.

DES0763N Die Aktualisierungsfrequenz ist nicht korrekt.

Erläuterung: Der angegebene Wert für die Aktualisierungsfrequenz ist ungültig.

Aktion: Geben Sie einen gültigen Wert für die Aktualisierungsfrequenz an.

DES0765N Die Datenbank ist für den DB2 Text Extender bereits aktiviert.

Erläuterung: Sie versuchen, eine Datenbank zu aktivieren, die bereits aktiviert ist.

Aktion: Entweder fahren Sie fort, ohne die Datenbank zu aktivieren, oder Sie verwenden DISABLE DATABASE, um die Datenbank zu inaktivieren, bevor Sie sie erneut aktivieren.

DES0766N Durch eine Aktion wurde die maximale Zeilenlänge der Tabelle oder einer temporären Tabelle überschritten.

Erläuterung: Der Befehl ENABLE TEXT COLUMN fügt eine Kennungsspalte zur Tabelle hinzu. Wenn die Tabelle bereits groß ist, kann dies dazu führen, daß die Zeilengröße der Tabelle den Maximalwert von 4005 überschreitet.

Der Befehl ENABLE TEXT COLUMN erstellt außerdem eine temporäre Tabelle, deren Größe proportional zur Anzahl der bereits aktivierten Textspalten ist. Wenn viele Textspalte bereits aktiviert sind, überschreitet die Größe der temporären Tabelle möglicherweise den Maximalwert.

Aktion: Verwenden Sie ENABLE TEXT COLUMN nur für Tabellen, die nicht dazu führen, daß dieses Limit überschritten wird.

DES0769W Warnung: Indexmerkmale wurden zwar angegeben, werden aber ignoriert. Tabelle *schema.tabelle* **ist eine Tabelle mit gemeinsam benutztem Index.**

Erläuterung: Die angegebene Tabelle ist eine Tabelle mit einem gemeinsam benutzten Index, daher können keine Indexmerkmale angegeben werden.

Aktion: Keine Aktion erforderlich.

DES0770N Die Umgebungsvariable *umgebungsvar* **ist nicht definiert.**

Erläuterung: Ein Parameter für einen Befehl wurde nicht angegeben und das System versuchte, den Standardwert aus der Umgebungsvariablen *umgebungsvar* zu lesen, aber diese Umgebungsvariable ist nicht definiert.

Aktion: Definieren Sie die erforderliche Umgebungsvariable.

DES0774N Die Länge des Wertes *wert* für die Variable *variable* ist nicht im gültigen Bereich.

Erläuterung: Die Wertlänge des Parameters *parameter* ist außerhalb des gültigen Bereichs.

Aktion: Geben Sie den Parameter *parameter* unter Verwendung einer gültigen Länge an.

DES0775N Der Wert für das Indexverzeichnis *wert* ist nicht korrekt.

Erläuterung: Der Wert für das Indexverzeichnis ist nicht korrekt, möglicherweise ist die Verzeichnislänge nicht korrekt.

Aktion: Geben Sie einen gültigen Wert für das Indexverzeichnis an.

DES0776N Der Tabellenbereichsname *name* ist nicht korrekt.

Erläuterung: Der angegebene Tabellenbereichsname ist nicht korrekt, möglicherweise ist die Länge des Tabellenbereichswertes nicht korrekt.

Aktion: Geben Sie einen gültigen Wert für den Tabellenbereich an.

DES0777N Der Tabellenbereich *tabellenbereich* ist im Datenbankverwaltungssystem nicht bekannt.

Erläuterung: Der angegebene Tabellenbereich ist dem Datenbanksystem nicht bekannt.

Aktion: Überprüfen Sie, daß der angegebene Tabellenbereich in der Datenbank existiert.

DES0778N *tabellenbereich* ist kein regulärer Tabellenbereich. Er wurde mit Hilfe des folgenden Schlüsselworts erstellt: *schlüsselwort*

Erläuterung: Der Datentyp für den angegebenen Tabellenbereich wird nicht unterstützt.

Aktion: Geben Sie einen regulären Tabellenbereich an.

DES0779I Die Indexierung wurde erfolgreich gestartet. Verwenden Sie 'GET INDEX STATUS', um den Indexierungsstatus zu überprüfen.

Erläuterung: Das Indexierungsprogramm wurde gestartet. Sie können den Befehl 'GET INDEX STATUS' verwenden, um den Status des Indexierungsprozesses zu überprüfen.

Aktion: Überprüfen Sie die Ausgabe des Befehls 'GET INDEX STATUS'.

DES0780I Die Reorganisation des Indexes wurde erfolgreich gestartet. Verwenden Sie 'GET INDEX STATUS', um den Indexstatus zu überprüfen.

Erläuterung: Das Reorganisationsprogramm wurde gestartet. Sie können den Befehl 'GET INDEX STATUS' verwenden, um den Status des Reorganisationsprozesses zu überprüfen.

Aktion: Überprüfen Sie die Ausgabe des Befehls 'GET INDEX STATUS'.

DES0789W Warnung: Die Partitionierungszuordnung der aktuellen Knotengruppe wurde aktualisiert, bitte das Dienstprogramm TXNCHECK aufrufen.

Erläuterung: Die Partitionierungszuordnung der aktuellen Knotengruppe wurde aktualisiert.

Aktion: Verwenden Sie den Befehl 'TXNCHECK'.

DES0800I Der Befehl *befehl* **wurde erfolgreich beendet.**

Erläuterung: Der angegebene Befehl wurde erfolgreich beendet.

Aktion: Keine Aktion erforderlich.

DES0810N Die abschließenden Anführungszeichen fehlen.

Erläuterung: Ein Anführungszeichen wurde gefunden, aber das zweite Anführungszeichen fehlt.

Aktion: Überprüfen Sie die Syntax des Befehls und wiederholen Sie den Versuch.

DES0811N *token* **wurde nicht erwartet. Überprüfen Sie die Indexmerkmale oder die Textdaten.**

Erläuterung: Die Indexmerkmale oder die Textdaten sind nicht korrekt.

Aktion: Überprüfen Sie die Syntax und wiederholen Sie den Befehl.

DES0812N Tabelle *schema.tabelle* **ist nicht vorhanden oder für den DB2 Text Extender nicht aktiviert.**

Erläuterung: Während der Ausführung des Befehls GET ist entweder der Name einer Datenbanktabelle nicht korrekt oder die Tabelle existiert nicht oder wurde noch nicht aktiviert.

Aktion: Wenn der Tabellenname korrekt ist, verwenden Sie GET STATUS, um zu überprüfen, daß die Tabelle aktiviert wurde. Aktivieren Sie die Tabelle und wiederholen Sie den Versuch.

DES0813N Tabelle *schema.tabelle* **ist nicht vorhanden, für DB2 Text Extender nicht aktiviert oder enthält keine Kennungsspalte** *kennungsspalte***.**

Erläuterung: Während der Ausführung des Befehls GET wurden keine Einträge für die Kennungsspalte in der Tabelle gefunden. Wenn die Tabelle existiert, ist sie nicht aktiviert oder sie enthält keine Kennungsspalte.

Aktion: Wenn der Tabellenname korrekt ist, verwenden Sie GET STATUS, um zu überprüfen, daß die Tabelle aktiviert wurde. Aktivieren Sie die Tabelle und wiederholen Sie den Versuch.

DES0814N Tabelle *schema.tabelle* **ist nicht vorhanden, für DB2 Text Extender nicht aktiviert oder in der Tabelle ist keine Textspalte aktiviert.**

Erläuterung: Die angegebene Tabelle ist nicht für den Text Extender aktiviert.

Aktion: Aktivieren Sie die Tabelle für den Text Extender.

DES0815N Leere Anführungszeichen "" gefunden. Innerhalb der Anführungszeichen muß ein Name stehen.

Erläuterung: Zwei aufeinanderfolgende Anführungszeichen wurden gefunden ohne Text dazwischen.

Aktion: Überprüfen Sie die Syntax und wiederholen Sie den Befehl.

DES0816N Das Wort *token* **wurde nicht erwartet. Eines der Schlüsselwörter** *schlüsselwort* **oder** *schlüsselwort* **angeben.**

Erläuterung: Ein unerwartetes Token wurde gefunden.

Aktion: Verwenden Sie im Befehl eines der Schlüsselwörter, die in der Nachricht angegeben sind.

DES0817N *token* **wurde nicht erwartet. Das Schlüsselwort** *schlüsselwort* **angeben.**

Erläuterung: Ein unerwartetes Token wurde gefunden.

Aktion: Verwenden Sie im Befehl das Schlüsselwort, das in der Nachricht angegeben ist.

DES0818N Unerwartetes Befehlsende. Das Schlüsselwort *schlüsselwort* **wird erwartet.**

Erläuterung: Ein Schlüsselwort fehlt.

Aktion: Verwenden Sie im Befehl das Schlüsselwort, das in der Nachricht angegeben ist.

DES0819N Unerwartetes Befehlsende. Eines der folgenden Schlüsselwörter wird erwartet: *schlüsselwort* **oder** *schlüsselwort***.**

Erläuterung: Ein Schlüsselwort fehlt.

Aktion: Verwenden Sie im Befehl eines der Schlüsselwörter, die in der Nachricht angegeben sind.

DES0820N Die Indexoption *index_option* **wird für den Indextyp** *index_typ* **nicht unterstützt.**

Erläuterung: Sie haben eine Indexoption angegeben, die für den angegebenen Indextyp nicht unterstützt wird.

Aktion: Unterstützte Indexoptionen für den angegebenen Indextyp finden Sie in der Dokumentation.

DES0821N Der Name *"token"* **ist zu lang. Für** *variable***-Namen sind nur** *nn* **Zeichen zulässig.**

Erläuterung: Ein Name ist zu lang.

Aktion: Geben Sie einen Namen in einer gültigen Länge an.

DES0822N Der Befehl enthält ein unbekanntes Token *"token"***. Das Befehlsende wurde erwartet.**

Erläuterung: Das Ende eines Befehls wurde gefunden, aber ein Schlüsselwort wird erwartet.

Aktion: Überprüfen Sie die Syntax des Befehls und wiederholen Sie den Versuch.

DES0823N Nach *"schema."* **wird ein Tabellenname erwartet.**

Erläuterung: Nach dem "." fehlt ein Tabellenname oder ein Funktionsname.

Aktion: Überprüfen Sie die Syntax des Befehls und wiederholen Sie den Versuch.

DES0824N Unerwartetes Befehlsende; ein *Schlüsselwort* **ist erforderlich.**

Erläuterung: Das Schlüsselwort in der Nachricht fehlt in der Syntax.

Aktion: Überprüfen Sie die Syntax des Befehls und wiederholen Sie den Versuch.

DES0825N *name* **ist kein bekannter Aliasname für eine Datenbank. Nur %d1 Zeichen sind zulässig.**

Erläuterung: Der angegebene Aliasname für die Datenbank ist dem Datenbanksystem unbekannt.

Aktion: Überprüfen Sie, daß der angegebene Aliasname gültig ist.

DES0826N Der Aliasname *alias* **für die Datenbank darf nicht in Anführungszeichen stehen.**

Erläuterung: Der Name in der Nachricht wurde als ein Aliasname für die Datenbank interpretiert. Er darf nicht in Anführungszeichen stehen.

Aktion: Überprüfen Sie die Syntax des Befehls und wiederholen Sie den Versuch.

DES0827N Ungültiger CCSID-Wert *"ccsid"***.**

Erläuterung: Die CCSID gehört nicht zu den vom Text Extender unterstützten CCSIDs.

Aktion: Eine Liste der unterstützten CCSIDs finden Sie in der Dokumentation.

DES0829N Der Benutzername *benutzer-ID* **darf nicht in Anführungszeichen stehen.**

Erläuterung: Sie haben einen Benutzernamen in Anführungszeichen eingegeben.

Aktion: Löschen Sie die Anführungszeichen.

DES0830N Parameter *"parameter"* **im Befehl** *enable/disable* **DATABASE wurde nicht erkannt. Das Befehlsende wurde erwartet.**

Erläuterung: Die Befehle ENABLE DATABASE und DISABLE DATABASE lassen keine Parameter zu.

Aktion: Geben Sie den Befehl erneut ein, ohne Parameter.

DES0831N Unerwartetes Befehlsende. Der Tabellenname fehlt.

Erläuterung: Für den Verwaltungsbefehl ist ein Tabellenname erforderlich.

Aktion: Geben Sie den entsprechenden Tabellennamen ein.

DES0832N Unerwartetes Befehlsende. Der Datenbankname fehlt.

Erläuterung: Für den Verwaltungsbefehl ist ein Datenbankname erforderlich.

Aktion: Geben Sie den entsprechenden Datenbanknamen ein.

DES0833N Unerwartetes Befehlsende. Der Spaltenname fehlt.

Erläuterung: Für den Verwaltungsbefehl ist ein Spaltenname erforderlich.

Aktion: Geben Sie den entsprechenden Spaltennamen ein.

DES0899N Unbekannter DB2TX-Befehl: *befehl***.**

Erläuterung: Der angegebene Befehl wird vom Text Extender nicht unterstützt.

Aktion: Geben Sie **db2tx ?** ein, um eine Liste der Befehle abzurufen.

DES0971N Das Indexverzeichnis kann einmal ohne Knoten bzw. mehrmals mit Knoten angegeben werden.

Erläuterung: Die Angabe des Indexverzeichnisses bzw. der Indexverzeichnisse ist nicht korrekt.

Aktion: Überprüfen Sie die Angabe des Indexverzeichnisses. Sie können ein Indexverzeichnis ohne Knotenangabe oder mehrere Indexverzeichnisse mit Knotenangabe angeben.

DES0972N Die Knotenangabe ist nicht korrekt. Eine Ziffer ohne Vorzeichen wird erwartet.

Erläuterung: Ein Wert, der keine Ziffer ist, wurde für die Knotennummer angegeben.

Aktion: Geben Sie eine Ziffer ohne Vorzeichen für die Knotennummer an.

DES0973N Die Knotenangabe ist nicht korrekt. Ein oder mehrere unerwartete Zeichen wurden vor der rechten runden Klammer gefunden.

Erläuterung: Die Knotenangabe ist syntaktisch nicht korrekt.

Aktion: Überprüfen Sie die Syntax der Knotenangabe und wiederholen Sie den Versuch.

DES0974N Die Reihenfolge von Knotennummern in einer TO-Klausel ist nicht korrekt (zweiter Knoten ist kleiner als erster Knoten).

Erläuterung: Die Knotenangabe ist syntaktisch nicht korrekt.

Aktion: Überprüfen Sie die Syntax der Knotenangabe und wiederholen Sie den Versuch.

DES0975N Die Syntax für die Angabe der Knotendaten ist nicht korrekt.

Erläuterung: Die Knotenangabe ist syntaktisch nicht korrekt.

Aktion: Überprüfen Sie die Syntax der Knotenangabe und wiederholen Sie den Versuch.

DES0976N Die Knotendaten sind nicht vollständig.

Erläuterung: Die Knotenangabe ist nicht vollständig - einige Informationen fehlen.

Aktion: Überprüfen Sie die Syntax der Knotenangabe und wiederholen Sie den Versuch.

DES0977N Die Knotendaten sind nicht vollständig. Die linke runde Klammer fehlt.

Erläuterung: Die Knotenangabe ist nicht vollständig - die linke runde Klammer fehlt.

Aktion: Korrigieren Sie die Knotenangabe und wiederholen Sie den Versuch.

DES0998N Die Länge des Dokumentmodellnamens ist nicht korrekt.

Erläuterung: Die Länge des Wertes für den Dokumentmodellnamen ist nicht korrekt.

Aktion: Überprüfen Sie den Wert für den Modellnamen und wiederholen Sie den Versuch.

DES0999N Die Syntax für die Angabe der Dokumentmodelle ist nicht korrekt.

Erläuterung: Die Angabe des Modellnamens (oder der Modellnamen) ist syntaktisch nicht korrekt.

Aktion: Überprüfen Sie die Syntax der Modellangabe und wiederholen Sie den Versuch.

DES9994N Ein Fehler in der Textsuchmaschine ist aufgetreten. Ursachencode: *ursachencode*

Erläuterung: In der Textsuchmaschine, die der Text Extender verwendet, ist ein Fehler aufgetreten.

Aktion: Überprüfen Sie den Ursachencode der Suchmaschine (verwenden Sie die TSE-Dokumentation). Wenn die aufgeführte Ursache nicht bei der Lösung des Problems hilft, finden Sie möglicherweise in der Datei desdiag.log oder in einer erstellten Trace-Datei zusätzliche Informationen. Hilft dies nicht, sammeln Sie die verfügbaren Informationen und rufen Sie den IBM Kundendienst an.

DES9995N Ein Fehler im Text Extender ist aufgetreten. Nachrichtentext: *nachrichtentext*

Erläuterung: Ein Fehler im Text Extender ist aufgetreten.

Aktion: Verwenden Sie die vom Text Extender gelieferte Nachricht, um das Problem zu lösen. Wenn die aufgeführte Nachricht nicht bei der Lösung des Problems hilft, finden Sie möglicherweise in der Datei desdiag.log oder in einer erstellten Trace-Datei zusätzliche Informationen. Hilft dies nicht, sammeln Sie die verfügbaren Informationen und rufen Sie den IBM Kundendienst an.

DES9996N Ein interner Fehler im DB2 Text Extender ist aufgetreten. Ursachencode: *ursachencode*

Erläuterung: Ein interner Verarbeitungsfehler ist aufgetreten.

Aktion: Überprüfen Sie, daß die Text Extender-Installation erfolgreich abgeschlossen wurde. Ist dies der Fall, notieren Sie den Ursachencode und rufen Sie den IBM Kundendienst an.

DES9997N SQL-Fehler aufgetreten. SqlState: *status* **SQL-Fehlercode:** *rc***; SqlErrorMessage:** *nachricht*

Erläuterung: SQL-Fehler aufgetreten.

Aktion: Gehen Sie anhand der SQL-Fehlernachricht vor, die mit der Nachricht angezeigt wird.

DES9998N SQL-Fehler aufgetreten. Es sind keine weiteren Informationen verfügbar.
DES9999N Keine zugehörige Fehlernachricht vorhanden.

Erläuterung: Ein interner Verarbeitungsfehler ist aufgetreten.

Aktion: Prüfen Sie die Diagnosenachricht, um das Problem zu lösen. Wenn der interne Fehler nicht durch einen Installationsfehler verursacht wurde, finden Sie möglicherweise in der Datei desdiag.log oder in einer erstellten Trace-Datei zusätzliche Informationen. Hilft dies nicht, sammeln Sie die verfügbaren Informationen und rufen Sie den IBM Kundendienst an.

Kapitel 14. Konfiguration

In diesem Kapitel werden die Text Extender-Umgebungsvariablen und -Konfigurationsdaten beschrieben. Mit beiden können Sie Standardwerte für viele Parameter angeben, die vom Text Extender benötigt werden.

Umgebungsvariablen

Die Umgebungsvariablen setzen die Standardwerte von Umgebungsparametern. Um die aktuelle Einstellung von Umgebungsvariablen anzuzeigen, verwenden Sie den Befehl GET ENVIRONMENT, der in „Die Einstellungen der Umgebungsvariablen anzeigen“ auf Seite 83 beschrieben wird.

DB2INSTANCE Der DB2 UDB-Exemplarname.

DB2DBDFT Name der Standarddatenbank. Der Name der DB2 UDB-Datenbank, die angenommen wird, wenn kein Datenbankname angegeben ist.

! **DB2TX_INSTOWNER** Der Text Extender-Exemplarname. Dies ist der Anmeldename
! des Benutzers, der Eigner des Exemplars ist. Er ist erforderlich, wenn ein
! Text Extender-Client eine Verbindung zu einem UNIX-Server herstellt.

DB2TX_INSTOWNERHOMEDIR Das Ausgangsverzeichnis des Exemplareigners (nur UNIX).

! **In einer Windows-Umgebung** sind die Text Extender-Umgebungsvariablen als
! Systemvariablen in der Seite für die Systemeigenschaften angegeben. Sie werden
! während der Installation zu Ihren Systemeinstellungen hinzugefügt.

! **In einer OS/2-Umgebung** werden die Text Extender-Umgebungsvariablen während der
! Installation zur Datei `CONFIG.SYS` hinzugefügt.

! **In einer UNIX-Umgebung** werden die Text Extender-Umgebungsvariablen (mit Aus-
! nahme von DB2DBDFT) in den folgenden Profil-Shell-Prozeduren angegeben:

! DB2TXCSHRC für die C-Shell

! DB2TXPROFILE für die Korn-Shell und die Bourne-Shell

! Kopieren Sie diese Dateien aus $DB2TXINSTOWNERHOMEDIR/DB2TX in Ihr Aus-
! gangsverzeichnis und passen Sie sie anschließend an Ihre Bedürfnisse an. Fügen Sie
! das Verzeichnis zu Ihrem Profil hinzu, so daß die Datenbankumgebung automatisch
! während des Anmeldens aufgebaut wird.

Konfiguration

! Sie benötigen außerdem die folgenden Umgebungsvariablen:

! **LANG** Gibt die Spracheinstellung an.

! **NLSPATH** Gibt den Pfad für die Text Extender-Nachrichtendatei an.

! **DB2ENVLIST** Liste von Umgebungsvariablen, die von DB2 an die UDFs übergeben
! werden.

! **RAHUSER** Gibt den Namen des Exemplareigners an, der für den DB2-Befehl `db2all`
! verwendet wird. (Nur in einer DB2 EEE-Umgebung.)

! Wenn Sie der Exemplareigner sind, werden diese Umgebungsvariablen automatisch für
! Sie gesetzt, wenn das Exemplar erstellt wird. Wenn Sie nicht der Exemplareigner sind,
! schließen Sie das DB2TX-Profil vom Exemplareigner ein, anstelle zu versuchen, diese
! Umgebungsvariablen manuell zu setzen.

Textkonfigurationseinstellungen

Textkonfigurationseinstellungen werden erstellt, wenn Sie eine Datenbank aktivieren. Dies sind Standardeinstellungen für Text-, Index- und Verarbeitungsmerkmale. Sie können diese Standardeinstellungen anzeigen und ändern. Siehe „Die Textkonfigurationseinstellungen anzeigen“ auf Seite 84 und „Die Textkonfiguration ändern“ auf Seite 50.

Textmerkmale

„Informationen zu Textdokumenten“ auf Seite 256 beschreibt die Dokumentformate, -sprachen und -CCSIDs, die vom Text Extender unterstützt werden. Standardwerte hierfür sind für verschiedene Verwaltungsbefehle erforderlich.

Wenn der Text Extender installiert ist, sind folgende Textkonfigurationseinstellungen als Standardwerte gesetzt:

FORMAT TDS

LANGUAGE Die Angabe LANGUAGE, die für die Datenbank gesetzt wurde

CCSID Die CCSID, die für die Datenbank gesetzt wurde

Indexmerkmale

DIRECTORY Verzeichnis, das zum Speichern des Indexes verwendet werden soll.

Anfangswert für UNIX-Systeme: *DB2TX_INSTOWNER*/db2tx/indexes.

Anfangswert für OS/2 und Windows NT: *DMBMMPATH*\instance*exemplarname* \db2tx\indexes, wobei *DMBMMPATH* der Pfad ist, in dem der Text Extender installiert ist. Für Windows-Clients muß DIRECTORY auf den Namen eines Verzeichnisses auf der Server-Maschine gesetzt werden.

INDEXTYPE Der zu verwendende Indextyp. In „Indextypen“ auf Seite 13 finden Sie dazu eine Beschreibung.

Anfangswert: LINGUISTIC

UPDATEFREQ Frequenz für die regelmäßige Indexaktualisierung. In „Die Frequenz von Indexaktualisierungen setzen“ auf Seite 262 finden Sie dazu eine Beschreibung.

Anfangswert: NONE

Verarbeitungsmerkmale

UPDATEINDEX Einstellung, mit der festgelegt wird, wann die erste Indexaktualisierung durchgeführt wird: entweder sofort, nachdem der Index erstellt ist (UPDATE), oder später gemäß den Einstellungen für die Aktualisierungsfrequenz (NOUPDATE).

Anfangswert: UPDATE

COMMITCOUNT Einstellung, mit der festgelegt wird, nach wie vielen Einfügungs- oder Aktualisierungsanweisungen der Text Extender eine DB2 UDB-Anweisung COMMIT ausgibt. Siehe „Eine Textspalte in einer großen Tabelle aktivieren“ auf Seite 63.

Anfangswert: 10000

Konfiguration

Informationen zu Textdokumenten

Jedes Textdokument, das Sie durchsuchen wollen, hat drei Merkmale, die wesentlich für den Text Extender sind:

Format

Sprache

CCSID (ID für codierten Zeichensatz)

Formate

Der Text Extender muß das Format (oder die Art) von Textdokumenten kennen, die Sie durchsuchen wollen, z.&rbl;B. WordPerfect oder ASCII. Diese Informationen werden beim Indexieren von Textdokumenten benötigt.

Folgende Arten von Textdokumenten werden unterstützt:

HTML Hypertext Markup Language

! **XML** Extended Markup Language

! **ASCII_SECTIONS** ASCII strukturiert, mit Abschnitten

TDS ASCII unstrukturiert

AMI AmiPro Architecture Version 4

FFT IBM Final Form Text: Document Content Architecture

MSWORD Microsoft Word, Versionen 5.0 und 5.5

RFT IBM Revisable Form Text: Document Content Architecture

RTF Microsoft Rich Text Format (RTF), Version 1

WP5 WordPerfect (OS/2 und Windows), Versionen 5.0, 5.1 und 5.2

Geben Sie für nicht unterstützte Dokumentarten eine numerische ID an. Gültige Werte sind 1 bis 100. Dieser Wert wird als Quellenformat an den Benutzer-Exit übergeben, der das ursprüngliche Format in TDS umwandelt.

Wenn während der Indexierung ein Dokument auftritt, dessen Dokumentart nicht einer der unterstützten Arten entspricht, bietet der Text Extender einen Exit, der das Dokument auf eine Platte schreibt und ein vom Benutzer zur Verfügung gestelltes Programm aufruft, um den Text in eines der unterstützten Formate zu extrahieren.

Konfiguration

Um den Benutzer-Exit zu aktivieren, editieren Sie die folgenden ASCII-Dateien:

```
Windows NT:
%DMBMMPATH%\instance\%DB2INSTANCE%\db2tx\descl.ini
%DMBMMPATH%\instance\%DB2INSTANCE%\db2tx\txinsnnn\dessrv.ini

UNIX:
$DB2TX_INSTOWNERHOMEDIR/db2tx/descl.ini
$DB2TX_INSTOWNERHOMEDIR/db2tx/txinsnnn/dessrv.ini
```

indem Sie die folgenden Anweisungen hinzufügen:

```
[DOCUMENTFORMAT]
USEREXIT=name_der_ausführbaren_datei
```

Dabei ist <name_der_ausführbaren_datei> der Name des Benutzer-Exit. Sie können einen vollständig qualifizierten Dateinamen angeben oder, wenn der Benutzer-Exit in einem Verzeichnis gespeichert ist, das in der Anweisung PATH angegeben ist, nur den Dateinamen.

Die Parameter des Benutzer-Exit müssen wie folgt angegeben werden:

```
<name_des_benutzer_exit>  -sourcefile   <quellendateiname>
                           -targetfile   <zieldateiname>
                           -sourceccsid  <quellen_ccsid>
                           -targetccsid  <ziel_ccsid>
                           -sourceformat <quellenformat>
                           -targetformat <zielformat>
```

Der Benutzer-Exit muß das Dokument aus <quellendateiname> lesen und das umgewandelte Dokument in <zieldateiname> schreiben. Die Dateinamen müssen vollständig qualifiziert sein. Die Zieldatei muß mit <ziel_ccsid> und <zielformat> übereinstimmen. Das Zielformat muß TDS sein. Die Ziel-CCSID muß 850 sein.

Während der Aktivierung muß ein anderes Format als TDS (ASCII unstrukturiert) als Format angegeben werden, um zu erzwingen, daß der Benutzer-Exit aufgerufen wird.

Konfiguration

Sprachen

Der Text Extender muß außerdem wissen, in welcher Sprache das Dokument geschrieben ist, so daß das korrekte Wörterverzeichnis für die Verarbeitung auf linguistischer Basis verwendet werden kann. Hier ist eine Liste der Sprachparameter, die Sie beim Aktivieren einer Textspalte oder externer Dokumente angeben können:

Portugiesisch (Brasilien)	BRAZILIAN
Französisch (Kanada)	CAN_FRENCH
Katalanisch	CATALAN
Vereinfachtes Chinesisch	S_CHINESE
Traditionelles Chinesisch	T_CHINESE
Dänisch	DANISH
Niederländisch	DUTCH
Finnisch	FINNISH
Französisch	FRENCH
Deutsch	GERMAN
Isländisch	ICELANDIC
Italienisch	ITALIAN
Japanisch	JAPANESE
Koreanisch	KOREAN
Norwegisch (Bokmal)	BM_NORWEGIAN
Norwegisch (Nynorsk)	NN_NORWEGIAN
Norwegisch (Bokmal und Nynorsk)	BMNN_NORWEGIAN
Portugiesisch	PORTUGUESE
Spanisch	SPANISH
Schwedisch	SWEDISH
Deutsch (Schweiz)	SWISS_GERMAN
Britisches Englisch	UK_ENGLISH
Amerikanisches Englisch	US_ENGLISH

CCSIDs

Jede DB2-Datenbank verwendet eine bestimmte Codepage für das Speichern von Zeichendaten. Der Text Extender als Anwendung, die mit DB2 arbeitet, läuft unter Verwendung derselben Codepage wie die Datenbank.

Dokumente können indexiert werden, wenn Sie in einer der folgenden CCSIDs vorliegen. Während der Suche wird die CCSID der Datenbank verwendet, um die CCSID des Suchbegriffs zu interpretieren.

Daten, die in DB2 UDB-Zeichendatentypen, z. B VARCHAR oder CLOB, gespeichert sind, werden durch DB2 UDB in die CCSID der Datenbank umgewandelt. Wenn Sie also eine Textspalte für die Suche aktivieren, verwenden Sie die CCSID der Datenbank als CCSID-Parameter. Wenn Sie eine Textspalte für die Suche aktivieren, können Sie die Datenumsetzung durch DB2 vermeiden, indem Sie den Datentyp BLOB oder einen binären Datentyp und die tatsächliche CCSID der Dokumente verwenden.

Anmerkung:

Die CCSIDs 861, 865 und 4946 werden nicht von DB2 UDB unterstützt. Um Dokumente, die diese CCSIDs haben, zu indexieren, speichern Sie die Dokumente in einer Spalte mit einem binären Datentyp (BLOB oder FOR BIT DATA).

EBCDIC

37	USA, Kanada
273	Österreich, Deutschland
277	Dänemark, Norwegen
278	Finnland, Schweden
280	Italien
284	Spanien, Lateinamerika
285	Großbritannien
297	Frankreich
437	USA
500	Internationales lateinisches Alphabet 1
871	Island
1027	Rußland

Konfiguration

ASCII

819 AIX, HP, SUN	Lateinisches Alphabet 1
850 AIX, OS/2	Lateinisches Alphabet 1
860 OS/2	Portugal
861 Siehe Anmerkung	Island
863 OS/2	Kanada
865 Siehe Anmerkung	Dänemark, Norwegen
866 OS/2	Rußland
915 AIX, OS/2, HP	Rußland
1250 WIN	Kroatien
1251 WIN	Rußland
1252 WIN	Lateinisches Alphabet 1

DBCS

932 AIX, OS/2	Japanisch, SBCS/DBCS kombiniert
942 OS/2	Japanisch, SBCS/DBCS kombiniert
943 OS/2, WIN	Japanisch, SBCS/DBCS kombiniert
5039 HP	Japanisch, SBCS/DBCS kombiniert
954 AIX, HP, SUN	Japanisch
949 OS/2	Koreanisch
970 AIX, HP, SUN	Koreanisch
1363 WIN	Koreanisch
948 OS/2	Chinesisch (traditionell), SBCS/DBCS kombiniert
950 AIX, HP, OS/2, SUN, WIN	Chinesisch (traditionell), SBCS/DBCS kombiniert
964 AIX, HP, SUN	Chinesisch (traditionell), SBCS/DBCS kombiniert
1381 OS/2, WIN	Chinesisch (vereinfacht), SBCS/DBCS kombiniert
1383 AIX, HP, SUN	Chinesisch (vereinfacht), SBCS/DBCS kombiniert
1386 AIX, OS/2, WIN	Chinesisch (vereinfacht), SBCS/DBCS kombiniert
4946 Siehe Anmerkung	Lateinische Alphabet 1 (CP850)
5039 HP	Japanisch

! **UNICODE**
! **UTF8**

! **UCS2**

! **Anmerkung:**

! Die CCSIDs 861, 865 und 4946 werden nicht von DB2 UDB unterstützt.
! Um Dokumente, die diese CCSIDs haben, zu indexieren, speichern Sie
! die Dokumente in einer Spalte mit einem binären Datentyp (BLOB oder
! FOR BIT DATA).

Die Frequenz von Indexaktualisierungen setzen

Wenn ein Textdokument zu einer Datenbank hinzugefügt wird oder wenn ein bestehendes Dokument in einer Datenbank geändert wird, muß das Dokument indexiert werden, damit der Inhalt des Indexes mit dem Inhalt der Datenbank synchron ist. Wenn ein Textdokument aus einer Datenbank gelöscht wird, müssen dessen Begriffe aus dem Index entfernt werden.

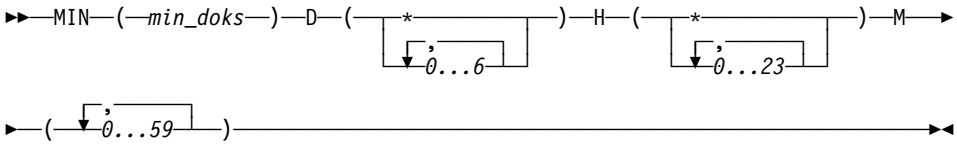
Informationen dazu, welche Dokumente neu, geändert oder gelöscht sind, werden automatisch durch Auslöser in einer Protokolltabelle gespeichert. Die in der Protokolltabelle aufgelisteten Dokumente werden das nächste Mal, wenn eine Indexaktualisierung stattfindet, indexiert.

Mit dem Befehl UPDATE INDEX können Sie einen Index sofort auf Anforderung aktualisieren, aber normalerweise aktualisieren Sie einen Index automatisch in Intervallen, die in der Umgebungsvariablen DB2TXUPDATEFREQ angegeben sind. Die Umgebungsvariable legt die Standardeinstellungen fest. Die Standardeinstellungen können überschrieben werden, wenn ein Index unter Verwendung des Befehls ENABLE TEXT COLUMN oder ENABLE TEXT TABLE erstellt wird. Die Aktualisierungsfrequenz kann für einen bestehenden Index geändert werden, indem der Befehl CHANGE INDEX SETTINGS verwendet wird.

Sie geben die Indexaktualisierungsfrequenz in bezug darauf an, wann die Aktualisierung durchgeführt werden soll und wie viele Textdokumente mindestens in der Warteschlange stehen müssen. Wenn zum bestimmten Zeitpunkt nicht genügend Dokumente in der Protokolltabelle stehen, wird der Index nicht aktualisiert.

Sie sollten das regelmäßige Indexieren sorgfältig planen; das Indexieren von Textdokumenten ist eine zeitaufwendige und ressourcenintensive Task. Die benötige Zeit hängt von vielen Faktoren ab, z.&rbl;B., wie viele Textdokumente seit der vorherigen Indexaktualisierung hinzugefügt oder geändert wurden, die Größe der Dokumente und die Leistungsfähigkeit des Prozessors.

Syntax



MIN min_doks Die Mindestanzahl an Textdokumenten, die in der Warteschlange stehen müssen, bevor der Index aktualisiert werden kann.

D Die Wochentage, an denen der Index aktualisiert wird:

- ***** Jeden Tag
- **0** Sonntag
- **1** Montag
- **2** Dienstag
- **3** Mittwoch
- **4** Donnerstag
- **5** Freitag
- **6** Samstag

H Die Stunden der angegebenen Tage, zu denen der Index aktualisiert wird:

- ***** Jede Stunde
- **0...23** Zu der angegebenen Stunde

M Die Minuten der angegebenen Stunde, zu denen der Index aktualisiert wird:

- **0...59** Zu der angegebenen Minute

Beispiel: `min(100) d(1,2,3,4,5) h(12,15) m(0)`

Wenn um 12:00 oder 15:00 Uhr von Montag bis Freitag mindestens 100 Textdokumente in der Warteschlange stehen, wird der Index aktualisiert.

Sie können mehrere Frequenzangaben kombinieren:

Beispiel: `min(1) d(*) h(22) m(0) ; min(100) d(1,2,3,4,5) h(12,15) m(0)`

Die Indexaktualisierung ist von Montag bis Freitag um 12:00 und 15:00 Uhr geplant wie zuvor, aber zusätzlich wird jeden Tag der Index um 22:00 Uhr aktualisiert, auch wenn nur ein Textdokument in der Protokolltabelle steht.

Konfiguration

Kapitel 15. Beispiel-API-Programm

Der Text Extender liefert ein Beispielprogramm, DESSAMP1.C, das sich im Verzeichnis SAMPLES befindet.

DESSAMP1.C ist ein Beispiel für ein Programm, das die Funktion DesGetSearchResultTable und Ihr Browser-Programm verwendet. Es folgt der Reihenfolge von API-Funktionsaufrufen, die in Abb. 15 auf Seite 119 gezeigt werden.

Sie benötigen Zugriff auf eine aktivierte Datenbank und eine aktivierte Textspalte. Führen Sie folgende Schritte aus, um dieses Programm auszuführen:

1. Wahlfrei. Kopieren Sie die Quellendatei DESSAMP1.C in ein lokales Verzeichnis auf Ihrer Client-Maschine.
2. Verwenden Sie die mitgelieferte Makefile (`DESSAMP.MAK` für OS/2 und Windows NT, `dessamp` für UNIX-Systeme), um die Beispieldateien zu kompilieren und zu verbinden.
3. Führen Sie das Dienstprogramm `DESRESTB` aus, um eine Ergebnistabelle in der Datenbank zu erstellen, die Sie mit dem Beispielcode verwenden wollen:

   `DESRESTB` *`datenbankname`*

   Diese Tabelle wird verwendet, um Informationen, wie z.&rbl;B. Suchergebnisse, zu speichern; sie hat die folgende Struktur:

Spalte	Datentyp
HANDLE	DB2TX.DB2TEXTH oder DB2TX.DB2TEXTFH
RANK	DOUBLE
MATCHES	INTEGER

Beispiel-API-Programm

Kapitel 16. Ursachencodes für Fehlerereignisse

In diesem Kapitel werden die Fehlerereignisse aufgelistet, die auftreten können, wenn der Text Extender Dokumente indexiert. Sie können beispielsweise in folgenden Fällen auftreten:

- Dokumente können nicht gefunden werden.
- Dokumente können nicht indexiert werden.
- Dokumente werden indexiert, aber ein Problem tritt auf.
- Ein Wörterverzeichnis für eine Sprache kann nicht gefunden werden.

Hinweis

Wenn ein Ursachencode nicht dokumentiert ist, führen Sie folgende Schritte aus:

1. Überprüfen Sie, daß ausreichend Plattenspeicherplatz verfügbar ist.
2. Sammeln Sie alle verfügbaren Fehlerinformationen:
   - Datei desdiag.log
   - Ereignisnachricht
3. Rufen Sie den IBM Kundendienst an.

1 Nicht genügend Speicher. Der Server hatte nicht mehr genügend Speicher zur Verfügung: verringern Sie die Arbeitsbelastung.

116 Syntaxfehler im Datenstrom.

280 Das Dokument wurde nicht indexiert. Eine der Indexdateien konnte nicht geöffnet werden.

281 Das Dokument wurde nicht indexiert. Eine der Indexdateien konnte nicht gelesen werden.

500 Das Dokument wurde nicht indexiert. Die Bibliotheksfunktionen konnten nicht geladen werden. Möglicherweise ist die DLL nicht verfügbar oder der Ressourcenpfad ist ungültig.

501 Das Dokument wurde nicht indexiert. Lib_Init in den Bibliotheksfunktionen für Systeme für unstrukturierte Dateien fehlgeschlagen: die DIT-Datei konnte nicht gefunden werden, sie befindet sich nicht in einem gültigen Verzeichnis oder der Inhalt der DIT-Datei ist nicht korrekt.

502 Das Dokument wurde nicht indexiert. Ein Fehler ist aufgetreten, während der Dokumentinhalt in der Bibliotheksfunktion LIB_read_doc_content gelesen wurde.

503 Das Dokument wurde nicht indexiert. Ein Fehler ist in der Bibliotheksfunktion LIB_access_doc aufgetreten.

Ursachencodes für Fehlerereignisse

504 Das Dokument wurde nicht indexiert. Die Bibliotheksfunktion LIB_doc_index_status gab einen Fehler zurück.

505 Das Schließen des Dokuments ist fehlgeschlagen. Die Bibliotheksfunktion LIB_close_doc gab einen Fehler zurück.

506 Das Beenden der Bibliotheksfunktionen ist fehlgeschlagen. Die Bibliotheksfunktion LIB_end gab einen Fehler zurück.

551, 553, 555, 556, 567, 569, 570, 571, 573, 574, 575, 576, 606, 608, 610, 619
Das Dokument wurde nicht indexiert. Eine der Indexdateien konnte nicht geöffnet werden.

624 Nicht genügend Speicher (alloc fehlgeschlagen). Der Server hatte nicht mehr genügend Speicher zur Verfügung: verringern Sie die Arbeitsbelastung.

659 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geöffnet werden.

660 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschrieben werden.

662 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geöffnet werden.

663 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschrieben werden.

665 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geöffnet werden.

667 Eine der während der Indexierung erstellten temporären Dateien konnte nicht geschrieben werden.

831 Das Dokument wurde nicht indexiert. Kein Text konnte gefunden werden. Die Dokumentlänge ist 0 Byte.

860 Fehler beim Öffnen der Datei (Wörterverzeichnis oder Thesaurus nicht gefunden).

1000 Ein Fehler ist während des Öffnens einer Datei aufgetreten. Bitte die Zugriffsberechtigungen prüfen.

1001 Ein Fehler ist während des Anhängens einer Datei aufgetreten. Bitte die Zugriffsberechtigungen prüfen.

1002 Ein Fehler ist während des Lesens einer Datei aufgetreten. Möglicherweise ist die Datei beschädigt.

1003 Ein Fehler ist während des Schreibens einer Datei aufgetreten. Bitte den Plattenspeicherplatz und die Zugriffsberechtigungen prüfen.

1007 Ein Fehler ist während des Erstellens einer Datei aufgetreten. Bitte die Zugriffsberechtigungen prüfen.

1010 Der angegebene Indexname ist bereits im Gebrauch. Einen anderen Indexnamen verwenden.

1011 Der angegebene Pfad ist bereits im Gebrauch. Einen anderen Standort verwenden.

1012 Derselbe Pfad wird für das Daten- und für das Arbeitsverzeichnis verwendet. Einen anderen Standort verwenden.

1013 Der angegebene Indexname ist ungültig. Indexnamen müssen in Großbuchstaben oder Ziffern angegeben werden und dürfen nicht länger als 8 Zeichen sein.

1017 Der Indexname ist unbekannt. Bitte die Schreibweise prüfen.

1020 Allgemeiner Dateifehler. Bitte die Zugriffsberechtigungen prüfen.

1072 Das Dokument wurde nicht indexiert. Die angegebene Codepage ist ungültig.

1073 Das Dokument wurde nicht indexiert. Die angegebene Codepage ist für diesen Indextyp ungültig.

1085 Das Dokument wurde nicht indexiert. Ein Fehler ist beim Lesen der Indexierungswarteschlange aufgetreten.

1086 Das Dokument wurde nicht indexiert. Die Indexierungswarteschlange ist leer.

1129 Kein Dokument wurde indexiert. Das Starten der Hintergrundverarbeitung ist fehlgeschlagen.

2000 Das Dokument wurde nicht indexiert. Die Dokumentart wird nicht unterstützt. Die Bibliotheksfunktion Lib_access_doc gab eine ungültige Dokumentart zurück.

2001 Das Dokument wurde nicht indexiert. Eine falsche Reihenfolge von Feldern wurde im Datenstrom des Dokuments festgestellt.

2002 Das Dokument wurde nicht indexiert. Ein falsch strukturiertes Feld wurde im Datenstrom des Dokuments festgestellt.

2003 Das Dokument wurde nicht indexiert. Nur ein Textkapitel ist für ein Dokument im Text Extender-Textformat zulässig.

2004 Das Dokument wurde nicht indexiert. Ein Hypermedia-Verweis, der länger als 512 Byte ist, wurde in einem Dokument im Text Extender-Textformat festgestellt.

2005 Das Dokument wurde nicht indexiert. Eine Sprache, die im Datenstrom des Dokuments angegeben ist, wird nicht unterstützt.

2006 Das Dokument wurde nicht indexiert. Eine CCSID, die im Datenstrom des Dokuments angegeben ist, wird nicht unterstützt.

Ursachencodes für Fehlerereignisse

2007 Das erwartete Dokumentformat, das von der Bibliothek oder von der Standardregel angegeben wird, ist nicht korrekt. Der Dokumentkopf ist für das Format nicht korrekt. Überprüfen Sie, ob die Standardregel ein Dokument mit einem speziellen Dokumentkopf ist und ändern Sie die Regel, wenn sie nicht korrekt ist.

2008 Das Dokument wurde nicht indexiert, da darauf nicht zugegriffen werden konnte.

2009 Das Dokument wurde nicht indexiert, da es im Gebrauch war und darauf nicht zugegriffen werden konnte.

2010 Das Dokument wurde nicht indexiert. Die angegebene CCSID ist nicht korrekt.

2011 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Das Seitenende muß das letzte Steuerelement im Hauptteil des Dokuments sein.

2012 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein strukturiertes Feld enthält eine falsche Längenangabe.

2013 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein falsches Steuerelement wurde im Dokument festgestellt.

2014 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein falsches Mehrbytesteuerelement oder strukturiertes Feld wurde im Dokument festgestellt.

2015 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Ein doppelter Dokumentparameter wurde gefunden.

2016 Das Dokument wurde nicht indexiert, da es kein gültiges IBM DCA RFT- oder FFT-Dokument ist. Eine leere Texteinheit wurde gefunden.

2018 Entweder ist das Dokument in einem Format, das nicht unterstützt wird, oder es steht ein “exclude”-Eintrag in der DIT-Datei für die Erweiterung des Dokuments. Überprüfen Sie, daß das Dokument eine Erweiterung hat, die eine Indexierung zuläßt.

2020 Das Dokument wurde nicht indexiert. Es ist weder ein WordPerfect-Dokument noch eine WordPerfect-Datei.

2021 Das Dokument wurde nicht indexiert. Es ist eine WordPerfect-Datei, aber kein WordPerfect-Dokument.

2022 Das Dokument wurde nicht indexiert. Diese Version von WordPerfect wird nicht unterstützt.

2023 Das Dokument wurde nicht indexiert. Es ist eine verschlüsselte WordPerfect-Datei. Speichern Sie das Dokument ohne Verschlüsselung.

2026 Ein END_TXT ist in einer Fußnote oder einer Endnote aufgetreten. Überprüfen Sie die WordPerfect-Datei, sie ist möglicherweise beschädigt.

2030 Das Dokument wurde nicht indexiert. Entweder ist es keine Microsoft Word-Datei oder es handelt sich um eine Version von Word, die nicht unterstützt wird.

2031 Das Dokument wurde nicht indexiert. Ein unerwartetes Dateiende wurde in einem Microsoft Word-Dokument festgestellt.

2032 Das Dokument wurde nicht indexiert. Ein falsches Steuerelement wurde in einem Microsoft Word-Dokument festgestellt.

2033 Das Dokument wurde nicht indexiert. Es wurde in *komplexem* Format mit der Option *fastsave* gesichert. Sichern Sie es mit ausgeschalteter Option *fastsave*.

2034 Das Dokument wurde nicht indexiert. Eine erforderliche Feldendemarke fehlt in einem Microsoft Word-Dokument.

2035 Das Dokument ist verschlüsselt. Speichern Sie das Dokument in Microsoft Word ohne Verschlüsselung.

2036 Dies ist ein Word for Macintosh-Dokument; es kann nicht verarbeitet werden. Speichern Sie das Dokument im Word for Windows-Format.

2037 Dieses Word-Dokument enthält eingebettete OLE-Objekte.

2040 Das Dokument wurde nicht indexiert, da es keine gültige ECTF-Datei ist.

2041 Das Dokument wurde nicht indexiert. Es enthält ein .SO LEN-Steuerelement, auf das keine Zahl folgt.

2042 Das Dokument wurde nicht indexiert. Es enthält ein .SO LEN-Steuerelement, auf das eine inkorrekte Zahl folgt. Es muß eine Zahl zwischen 1 und 79 sein.

2043 Das Dokument wurde nicht indexiert. Nur ein .SO DOC-Steuerelement ist zulässig. Sichern Sie jedes ECTF-Dokument jeweils in einer separaten Datei.

2044 Das Dokument wurde nicht indexiert. Auf ein .SO HDE-Steuerelement muß ein Anfangsbefehl und ein Endbefehl folgen.

2045 Das Dokument wurde nicht indexiert. Der Anfangs- oder Endbefehl für den Hypermedia-Verweis ist zu lang.

2046 Das Dokument wurde nicht indexiert. Das Dokument enthält Text vor dem .SO DOC-Steuerelement.

2047 Das Dokument wurde nicht indexiert. Das Dokument enthält Text vor einem .SO PID-Steuerelement.

2048 Das Dokument wurde nicht indexiert. Ein Endbefehl fehlt nach einem Anfangsbefehl.

2049 Das Dokument wurde nicht indexiert. Ein Hypermedia-Verweis, der länger als 80 Byte ist, wurde festgestellt.

Ursachencodes für Fehlerereignisse

2050 Das Dokument wurde nicht indexiert. Falsche Befehle wurde nach einem .SO HDE-Steuerelement festgestellt.

2051 Das Dokument wurde nicht indexiert. Ein Zeilenende wurde nach einem .SO-Steuerelement festgestellt.

2052 Das Dokument wurde nicht indexiert. Ein unerwartetes Textende wurde festgestellt.

2060 Das Dokument wurde nicht indexiert. Entweder ist es kein AmiPro-Dokument oder es handelt sich um eine Version von AmiPro, die nicht unterstützt wird.

2061 Das Dokument wurde nicht indexiert. Die Länge eines Steuerelements in einem AmiPro-Dokument ist zu lang.

2062 Das Dokument wurde nicht indexiert. Diese Version von AmiPro wird nicht unterstützt. Nur AmiPro Architecture Version 4 wird unterstützt.

2063 AmiPro Style Sheets wurden nicht indexiert.

2064 Das Dokument wurde nicht indexiert. Ein falscher Zeichensatz wurde festgestellt. Nur Lotus Character Set 82 (Windows ANSI) wird unterstützt.

2065 Das Dokument wurde nicht indexiert. Ein unerwartetes Dateiende wurde in einem AmiPro-Dokument festgestellt.

2072 Das Dokument kann nicht durchsucht werden, da es verschlüsselt ist.

2073 Das Dokumentformat ist nicht konsistent.

2074 Für das Dokument ist das Markierungsbit “bad file” gesetzt.

2080 Das Dokument wurde nicht indexiert. Entweder ist es kein RTF-Dokument oder es handelt sich um eine Version von RTF, die nicht unterstützt wird.

2081 Das Dokument wurde nicht indexiert. Ein RTF-Steuerwort, das zu lang ist, wurde festgestellt.

2083 Das Dokument wurde nicht indexiert. Die Macintosh-Codepage wird nicht unterstützt.

2084 Das Dokument wurde nicht indexiert. Es ist ein RTF-Dokument, aber diese Version von RTF wird nicht unterstützt. Nur RTF Version 1 wird unterstützt.

2100 Das Dokument ist beschädigt oder aus einem anderen Grund nicht lesbar. Ein neuer allgemeiner Parser könnte das Problem beheben.

2101 Das Dokument kann nicht indexiert werden, da es leer ist oder keinen Text enthält. Überprüfen Sie, ob das Dokument nur Grafiken enthält.

2102 Das Dokument kann nicht indexiert werden, da es entweder kennwortgeschützt oder verschlüsselt ist.

2105 Die Dokumentart ist bekannt, aber der Filter ist nicht verfügbar.

2106 Das Dokument kann nicht indexiert werden, da es leer ist.

2107 Das Dokument kann nicht indexiert werden, da es nicht geöffnet werden kann. Den Dokumentzugriff prüfen.

2112 Das Dokument kann nicht indexiert werden, da es eine ausführbare Datei ist.

2113 Das Dokument kann nicht indexiert werden, da es komprimiert ist.

2114 Das Dokument kann nicht indexiert werden, da es eine Grafik ist. Wenn das Grafikdokumentformat ein akzeptables Textstück zurückgibt, fordern Sie an, dieses Dokumentformat in den Indexierungsprozeß einzuschließen.

2120 Die Ausgabedatei des Benutzer-Exit existiert nicht oder es kann nicht darauf zugegriffen werden. Eine neue allgemeine Parser-Version könnte das Problem beheben.

2121 Die Ausgabedatei kann nicht zum Lesen geöffnet werden oder sie ist leer. Eine neue allgemeine Parser-Version könnte das Problem beheben.

2122 Es wurde versucht, eine Ausgabedatei des Benutzer-Exit zu verwenden, aber es wurde kein Dateiname angegeben oder im Objekt gesetzt.

2130 Das Benutzer-Exit-Programm konnte nicht ausgeführt werden. Überprüfen Sie, ob sich die ausführbare Datei in dem Pfad befindet, der durch die Umgebungsvariable PATH gesetzt ist. Erstellen Sie einen Trace und einen Speicherauszug, um zusätzliche Informationen zu den Umgebungsrückkehrcodes (errno) zu erhalten.

2131 Das Benutzer-Exit-Programm ist mit einem fehlerhaften Rückkehrcode fehlgeschlagen. Erstellen Sie einen Trace und einen Speicherauszug, um zusätzliche Informationen zu den Umgebungsrückkehrcodes (errno) zu erhalten.

Ursachencodes für Fehlerereignisse

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, daß nur Programme, Produkte oder Dienstleistungen von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Dienstleistungen können auch andere ihnen äquivalente Produkte, Programme oder Dienstleistungen verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte der IBM verletzen. Die Verantwortung für den Betrieb von Fremdprodukten, Fremdprogrammen und Fremdservices liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanfragen sind schriftlich an IBM Europe, Director of Licensing, 92066 Paris La Defense Cedex, France, zu richten. Anfragen an obige Adresse müssen auf englisch formuliert werden.

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. Die Angaben in diesem Handbuch werden in regelmäßigen Zeitabständen aktualisiert. Die Änderungen werden in Überarbeitungen oder in Technical News Letters (TNLs) bekanntgegeben. IBM kann jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängigen, erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Corporation
J74/G4
555 Bailey Avenue
P.O. Box 49023
San Jose, CA 95161-9023
U.S.A.

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des im Handbuch aufgeführten Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt im Rahmen der Allgemeinen Geschäftsbedingungen der IBM, der Internationalen Nutzungsbedingungen der IBM oder einer äquivalenten Vereinbarung.

Diese Informationen enthalten Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufes. Sie sollen nur die Funktionen des Lizenzprogrammes illustrieren; sie können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

COPYRIGHT-LIZENZ

Diese Veröffentlichung enthält Beispielanwendungsprogramme, die in Quellensprache geschrieben sind. Sie dürfen diese Beispielprogramme kostenlos kopieren, ändern und verteilen, wenn dies zu dem Zweck geschieht, Anwendungsprogramme zu entwickeln, verwenden, vermarkten oder zu verteilen, die mit der Anwendungsprogrammierschnittstelle konform sind, für die diese Beispielprogramme geschrieben werden. Diese Beispielprogramme wurden nicht unter allen denkbaren Bedingungen getestet.

Wird dieses Buch als Softcopy (Book) angezeigt, erscheinen möglicherweise keine Fotografien oder Farbabbildungen.

Informationen zur Programmierschnittstelle

Diese Veröffentlichung soll Sie bei der Verwaltung des Text Extender und bei der Entwicklung von Programmen zur Verwendung mit dem Text Extender unterstützen. Diese Veröffentlichung dokumentiert die allgemeine Programmierschnittstelle und zugehörige Informationen, die durch den Text Extender zur Verfügung gestellt werden.

Allgemeine Programmierschnittstellen ermöglichen Ihnen, Programme zu schreiben, die die Services des Text Extender anfordern. Sie können die Laufzeitfunktion des Text Extender, die für die Anwendung, die Sie entwickeln, erforderlich ist, auf die Client- oder Server-Maschinen kopieren. Um die Laufzeitfunktion auf einer Client-Maschine zu installieren, lesen Sie die Installationsanweisungen, die in der Datei README.TXT für Ihr Betriebssystem auf der CD-ROM DB2 Extender-Clients zur Verfügung stehen.

Marken

Folgende Namen sind in gewissen Ländern Marken oder Dienstleistungsmarken der IBM Corporation:

AIX	DB2 Universal Database	PS/2
DB2	IBM	
DB2 Extender	OS/2	

Microsoft, Windows, Windows NT und das Logo von Windows sind Marken der Microsoft Corporation.

UNIX ist eine eingetragene Marke und wird ausschließlich von der X/Open Company Limited lizenziert.

Intel ist eine eingetragene Marke von Intel.

Andere Namen von Unternehmen, Produkten und Dienstleistungen können Marken oder Dienstleistungsmarken anderer Unternehmen sein.

Glossar

A

Abrufen. Das Finden eines Textdokuments unter Verwendung eines Sucharguments in einer der Suchfunktionen des Text Extender.

Aktivieren. Eine Datenbank, eine Texttabelle oder eine Textspalte für die Verwendung durch den Text Extender vorbereiten.

Aktualisierungsfrequenz. Die Frequenz, mit der ein Textindex aktualisiert wird, ausgedrückt in Tag, Stunden und Minuten und in der Mindestanzahl an Dokumentnamen, die in der *Protokolltabelle* für die Indexierung aufgelistet sein müssen, bevor die Indexierung stattfinden kann.

Anwendungsprogrammierschnittstelle (API). Eine allgemein einsetzbare Schnittstelle zwischen Anwendungsprogrammen und den Informationsabfrageservices des Text Extender.

Anzeigen. Text auf einem Computerbildschirm sichtbar machen.

API. Anwendungsprogrammierschnittstelle (Application Programming Interface).

Auslöser. Ein Mechanismus, der immer automatisch Informationen über Dokumente, die indexiert werden müssen, zu einer *Protokolltabelle* hinzufügt, wenn ein Dokument in einer Textspalte hinzugefügt, geändert oder gelöscht wird.

B

Befehlszeilenprozessor. Ein Programm mit dem Namen DB2TX, das

Ihnen ermöglicht, Text Extender-Befehle einzugeben.

die Befehle verarbeitet.

das Ergebnis anzeigt.

Benutzerdefinierte Funktion (UDF). Eine SQL-Funktion, die von einem DB2-Benutzer erstellt wurde, im Gegensatz zu einer SQL-Funktion, die von DB2 zur Verfügung gestellt wird. Der Text Extender stellt Suchfunktionen, z.&rbl;B. CONTAINS, in Form von UDFs zur Verfügung.

Benutzerdefinierter eindeutiger Typ (UDT). Ein Datentyp, der von einem DB2-Benutzer erstellt wurde, im Gegensatz zu einem Datentyp, der von DB2 zur Verfügung gestellt wird, z.&rbl;B. LONG VARCHAR.

Boolesche Suche. Eine Suche, in der ein oder mehrere Suchbegriffe unter Verwendung von Booleschen Operatoren kombiniert werden.

BOUND-Suche. Eine Suche in koreanischen Dokumenten, bei der Wortgrenzen beachtet werden.

Browser. Eine Text Extender-Funktion, mit der Sie Text auf einem Computerbildschirm anzeigen können.

C

CCSID. ID des codierten Zeichensatzes (Coded Character Set Identifier).

Codepage. Eine Zuordnung von Schriftzeichen und Steuerfunktionsbedeutungen für alle Codepunkte. Beispielsweise die Zuordnung von Zeichen und Bedeutungen für 256 Codepunkte für einen 8-Bit-Code.

COUNT. Ein Schlüsselwort, mit dem die Anzahl der Ebenen (die Tiefe) von Begriffen im Thesaurus angegeben wird, die verwendet werden sollen, um den Suchbegriff für eine bestimmte Relation zu erweitern.

D

Dateikennung. Siehe *Kennung*.

Datenstrom. Informationen, die von einer API-Funktion zurückgegeben werden, die Text (mindestens einen Absatz), der den gesuchten Begriff enthält, und Informationen zur Hervorhebung des gefundenen Begriffs in diesem Text umfassen.

DBCS. Unterstützung von Doppelbytezeichensätzen.

DB2 Extender. Eines aus einer Gruppe von Programmen, mit denen Sie Datentypen speichern und abrufen, die über die traditionellen numerischen und Zeichendaten hinausgehen, z.B. Abbild-, Audio- und Videodaten sowie komplexe Dokumente.

Dokument. Siehe *Textdokument*.

Dokumentkennung. Siehe *Kennung*.

! **Dokumentmodell**. Die Definition der Struktur eines
! Dokuments in bezug auf die Abschnitte, die es enthält.
! Ein Dokumentmodell verdeutlicht dem Text Extender
! beim Indexieren die Abschnitte innerhalb von Doku-
! menten. Ein Dokumentmodell listet die
! Formatierungssteuerzeichen auf, die die Abschnitte
! kennzeichnen. Für jedes Steuerzeichen können Sie
! einen beschreibenden Abschnittsnamen angeben, der
! bei Abfragen für diesen Abschnitt verwendet wird. Sie
! können ein oder mehrere Dokumentmodelle in einer
! Dokumenthodelldatei angeben.

Dualer Index. Ein *Textindex*, der die Merkmale eines *präzisen Indexes* und eines *linguistischen Indexes* besitzt. Siehe auch *Ngram-Index*.

E

Eindeutiger Typ. Siehe *Benutzerdefinierter eindeutiger Typ*.

Erweitern. Die Aktion, bei der zusätzliche Begriffe zu einem Suchbegriff hinzugefügt werden, die aus einem Thesaurus stammen.

Erweiterte Übereinstimmung. Ein Prozeß, bei dem ein *Wörterverzeichnis* verwendet wird, um Begriffe hervorzuheben, die nicht offensichtlich mit dem Suchbegriff übereinstimmen.

Escape-Zeichen. Ein Zeichen, das angibt, daß das nachfolgende Zeichen nicht als *Platzhalterzeichen* interpretiert werden soll.

Exemplar. Eine logische Text Extender-Umgebung. Sie können mehrere Exemplare des Text Extender auf derselben Workstation haben, aber nur jeweils ein Exemplar für jedes DB2-Exemplar. Sie können diese Exemplare zu folgenden Zwecken verwenden:

Trennen der Entwicklungsumgebung von der Produktionsumgebung

Einschränken von sensiblen Daten auf eine bestimmte Personengruppe

Exemplarvariable. Eine Variable, mit der ein Standardwert für den Namen des *Exemplar*eigners oder der Name des Ausgangsverzeichnisses des Exemplareigners geliefert wird.

Extender. Siehe *DB2 Extender*.

Externe Datei. Ein Textdokument in Form einer Datei, die im Dateisystem des Betriebssystem gespeichert ist, und nicht in Form einer Zelle in einer Tabelle, die von DB2 gesteuert wird.

F

Feinanpassung. Das Hinzufügen von Suchkriterien aus einer vorherigen Suche zu anderen Suchkriterien, um die Anzahl der *Übereinstimmungen* zu verringern.

Format. Die Art eines Dokument, z.&rbl;B. ASCII oder WordPerfect.

Freitextsuche. Eine Suche, in der der Suchbegriff als unformatierter Text ausgedrückt wird – ein Ausdruck oder ein Satz, der in natürlicher Sprache das Subjekt, nach dem gesucht werden soll, beschreibt.

Funktion. Siehe *Zugriffsfunktion*.

H

Hervorhebungsinformationen. Siehe *Datenstrom*.

Hybridsuche. Eine Kombination aus *Booleschen Suche* und *Freitextsuche*.

I

Inaktivieren. Eine Datenbank, eine Texttabelle oder eine Textspalte in den Zustand zurückbringen, den sie vor der Aktivierung für den Text Extender hatte, indem die Elemente, die während des Aktivierungsprozesses erstellt wurden, gelöscht werden.

Indexieren. Das Extrahieren von wesentlichen Begriffen aus Text und das Speichern dieser Begriffe in einem *Textindex*.

Indexmerkmale. Die Merkmale eines *Textindexes* bestimmen:

Das Verzeichnis, in dem der Index gespeichert ist

Den Indextyp

Die Frequenz, mit der der Index aktualisiert wird

Wann die erste Indexaktualisierung stattfinden soll

Indextyp. Ein Merkmal eines *Textindexes*, das festlegt, ob der Index exakte oder linguistische Formen von Dokumentbegriffen enthält oder beide. Siehe *Präziser Index*, *Linguistischer Index*, *Dualer Index* und *Ngram-Index*.

Initialisierte Kennung. Eine *Kennung*, die vorab vorbereitet wurde, die nur das Textformat oder die Textsprache oder beides enthält.

K

Katalogsicht. Eine Sicht einer Systemtabelle, die vom Text Extender zu Verwaltungszwecken erstellt wurde. Eine Katalogsicht enthält Informationen zu den Tabellen und Spalten, die zur Verwendung durch den Text Extender aktiviert worden sind.

Kennung. Ein Binärwert, der ein Textdokument identifiziert. Dazu gehört:

Eine Dokument-ID

Name und Standort des zugeordneten Indexes

Die *Textdaten* des Dokuments

Wenn sich das Dokument in einer externen Datei befindet, die nicht von DB2 gesteuert wird, Pfad und Name der Datei

Eine Kennung wird für jedes Textdokument in einer Textspalte erstellt, wenn diese Spalte für die Verwendung durch den Text Extender *aktiviert* ist.

Knoten. Ein Server in einer Umgebung für *partitionierte Datenbanken*. Siehe auch *Logischer Knoten*, *Physischer Knoten* und *Knotengruppe*.

Knotengruppe. Eine benannte Untergruppe von einem oder mehreren Datenbankpartitions-Servern. Ein *Knoten*, der einer phyisch separaten Maschine zugeordnet ist. Siehe auch *Logischer Knoten*.

L

Linguistischer Index. Ein *Textindex*, der Begriffe enthält, die durch die Verarbeitung auf linguistischer Basis auf ihre Grundform reduziert wurden. “Mäuse” würde beispielsweise als “maus” indexiert. Siehe auch *Präziser Index*, *Ngram-Index* und *Dualer Index*.

Logischer Knoten. Ein *Knoten*, der mit anderen Knoten derselben physischen Maschine zugeordnet ist. Siehe auch *Physischer Knoten*.

M

Mehrindextabelle. Eine DB2-Tabelle, deren Textspalten individuelle *Textindizes* haben. Siehe auch *Tabelle mit einem gemeinsam benutzten Index*.

Merkmalsuche. Ein Suche nach Begriffen, z.&rbl;B. Eigennamen von Personen, Orten und Organisationen, die in einem linguistischen Index durchgeführt wird, der mit der Indexierungsoption FEATURE_EXTRACTION erstellt wurde.

N

Ngram-Index. Ein *Textindex*, der DBCS-Dokumente und die Suche nach grober Übereinstimmung in SBCS-Dokumenten unterstützt. Siehe auch *Linguistischer Index*, *Präziser Index* und *Dualer Index*.

P

Partitionierte Datenbank. Eine Datenbank, die aus mehreren Teilen besteht, von denen jedes durch einen separaten Datenbankpartitions-Server verwaltet wird.

Physischer Knoten. Ein *Knoten*, der einer phyisch separaten Maschine zugeordnet ist. Siehe auch *Logischer Knoten*.

Platzhalterzeichen. Ein Zeichen, das zum Repräsentieren eines wahlfreien Zeichens am Anfang, in der Mitte und am Ende eines Suchbegriffs verwendet wird. Platzhalterzeichen werden normalerweise verwendet, um Variationen eines Begriffs in einem präzisen Index zu finden.

Präziser Index. Ein *Textindex*, der Begriffe exakt so enthält, wie sie im Textdokument auftreten, aus dem sie extrahiert wurden. Siehe auch *Linguistischer Index*, *Ngram-Index* und *Dualer Index*.

Profil. Siehe *Umgebungsprofil*.

Protokolltabelle. Eine vom Text Extender erstellte Tabelle, die Informationen dazu enthält, welche Textdokumente indexiert werden sollen. *Auslöser* werden verwendet, um diese Informationen immer in einer Protokolltabelle zu speichern, wenn ein Dokument in einer aktivierten Textspalte hinzugefügt, geändert oder gelöscht wird.

R

Rangordnung. Ein absoluter Wert vom Typ DOUBLE zwischen 0 und 1, der angibt, wie gut das Dokument in Relation zu den anderen gefundenen Dokumenten mit den Suchkriterien übereinstimmt. Der Wert gibt die Anzahl der im Dokument gefundenen Übereinstimmungen in Relation zur Größe des Dokuments an.

Regelmäßiges Indexieren. Indexieren zu vorab festgelegten Zeitintervallen, ausgedrückt in Tag, Stunden und Minuten und in der Mindestanzahl an Dokumentnamen, die in der *Protokolltabelle* für die Indexierung aufgelistet sein müssen, bevor die Indexierung stattfinden kann.

S

SBCS. Unterstützung von Einzelbytezeichensätzen.

Shell-Profil. Siehe *Umgebungsprofil*.

Sprache. Der Name eines *Wörterverzeichnisses*, das beim *Indexieren*, Suchen und *Anzeigen* verwendet werden soll.

Stoppwort. Ein allgemeines Wort, z.&rbl;B. “bevor”, in einem *Textdokument*, das aus dem *Textindex* ausgeschlossen werden soll und ignoriert wird, wenn es in einem *Suchargument* angegeben wird.

Suchargument. Die Bedingungen, die beim Ausführen einer Suche angegeben werden, die aus einem oder mehreren Suchbegriffen sowie Suchparametern bestehen.

Suche nach grober Übereinstimmung. Eine Suche, bei der Wörter gefunden werden können, deren Schreibweise der des Suchbegriffs ähnelt.

T

Tabelle mit einem gemeinsam benutzten Index. Eine DB2-Tabelle, deren Textspalten einen Textindex gemeinsam benutzen. Siehe auch *Mehrindextabelle*.

Textdaten. Merkmale eines *Textdokuments*, die folgendes beschreiben:

Die *CCSID*

Das *Format*

Die *Sprache*

Textdokument. Text vom Typ CHAR, GRAPHIC, VARGRAPHIC, LONG VARGRAPHIC, DBCLOB, VARCHAR, LONG VARCHAR oder CLOB, der in einer DB2-Tabelle gespeichert ist.

Textindex. Eine Sammlung von wesentlichen Begriffen, die aus Textdokumenten extrahiert wurden. Jeder Begriff ist dem Dokument zugeordnet, aus dem er extrahiert wurde. Eine wesentliche Verbesserung der Suchzeit wird erzielt, wenn der Index und nicht die Dokumente selbst durchsucht werden. Siehe auch *Präziser Index*, *Linguistischer Index* und *Dualer Index*.

Textkonfiguration. Standardeinstellungen für Index-, Text- und Verarbeitungswerte.

Textspalte. Eine Spalte, die *Textdokumente* enthält.

Texttabelle. Eine DB2-Tabelle, die *Textspalten* enthält.

Trace-Funktion ausführen. Die Aktion, bei der Informationen in einer Datei gespeichert werden, die zu einem späteren Zeitpunkt verwendet werden kann, um die Ursache eine Fehlers zu finden.

U

UDF. Benutzerdefinierte Funktion.

UDT. Benutzerdefinierter eindeutiger Typ.

Übereinstimmung. Das Vorkommen eines Suchbegriffs in einem Textdokument.

Umgebungsprofil. Eine mit dem Text Extender gelieferte Prozedur, die Einstellungen für *Umgebungsvariablen* enthält.

Umgebungsvariable. Eine Variable, mit der Standardwerte für die Werte für die Text Extender-Umgebung geliefert werden.

V

Verwaltung. Die Task, Textdokumente für die Suche vorzubereiten, Indizes zu verwalten und Statusinformationen abzurufen.

Vorkommen. Synonym zu *Übereinstimmung*.

W

Wörterverzeichnis. Eine Sammlung von sprachbezogenen linguistischen Informationen, die der Text Extender während der Textanalyse, der Indexierung, der Abfrage und der Hervorhebung von Dokumenten in einer bestimmten Sprache verwendet.

Z

Zugriffsfunktion. Eine vom Benutzer gestellte Funktion, die den Datentyp von Text, der in einer Spalte gespeichert ist, in einen Typ umsetzt, der vom Text Extender verarbeitet werden kann.

Index

Sonderzeichen

| (OR), Operator in Suchargumenten
  Suchargumentsyntax 197
  Verwendung 103
& (AND), Operator in Suchargumenten
  Suchargumentsyntax 197
  Verwendung 103

A

Abfrage, Verarbeitung auf linguistischer Basis 24
Abkürzungen
  Editieren einer Abkürzungsdatei 52
  Listen 31
Abrufen der Suchergebnistabelle, API-Funktion
  Beschreibung 216
  Verwendung 120
Abrufen von Anzeigeinformationen, API-Funktion
  Beschreibung 208
  Verwendung 121
Abrufen von Übereinstimmungen, API-Funktion
  Beschreibung 211
  Verwendung 122
Abschnitte in Dokumenten
  Aktivieren der Abschnittsunterstützung 68
  MODEL, Schlüsselwort in der Suchsyntax 195
  SECTION, Schlüsselwort in der Suchsyntax 195
  Suchbeispiel 106
Aktualisieren eines Textindexes
  Ändern der Frequenz 79
  regelmäßig 262
  UPDATEFREQ in CHANGE INDEX SETTINGS 130
  UPDATEFREQ in CHANGE TEXT CONFIGURATION 132
Aktualisierungsfrequenz
  ändern 79
  Beschreibung 262
  GET INDEX SETTINGS, Befehl 154
  Standardwert in den Textkonfigurationseinstellungen 255
  Syntax 262
  UPDATEFREQ in CHANGE INDEX SETTINGS 130
  UPDATEFREQ in CHANGE TEXT CONFIGURATION 132
Aktualisierungsstatus, anzeigen
  Beispiel und Ausgabe 84
  Syntax 156
Aktualisierungsstatus, zurücksetzen
  Beispiel 79
  Syntax 160
AmiPro, Dokumentart 256
Analyse von Text
  für die Anzeige 27
  für die Indexierung 21
AND
  Boolescher Operator 103
  Schlüsselwort in einem Suchargument 197
Anwendungsprogrammierschnittstelle (API)
  Abrufen des Zeigers für Hervorhebungsinformationen 211
  Abrufen einer Suchergebnistabelle 120
  Abrufen von Anzeigeinformationen 121
  Abrufen von Übereinstimmungen 122
  Anzeigefunktionen 117
  Beenden einer Anzeigesitzung 123
  DesCloseDocument, Funktion 206
  DesEndBrowseSession, Funktion 207
  DesFreeBrowseInfo, Funktion 208
  DesGetBrowseInfo, Funktion 208
  DesGetMatches, Funktion 211
  DesGetSearchResultTable, Funktion 216
  DesOpenDocument, Funktion 220
  DesStartBrowseSession, Funktion 222
  Freigeben des Speichers für Anzeigeinformationen 123
  Hervorhebungsinformationen 211
  Nachrichten 235
  Öffnen eines Dokuments für die Anzeige 122
  Programmbeispiel 265
  Referenz 205
  Rückkehrcodes 225
  Schließen eines Dokuments 123
  Starten einer Anzeigesitzung 121
  Suchen nach Text 120
  Suchfunktionen 117
  Überblick 118
  Zusammenfassung 205
Anzeigefunktionen 117
Anzeigen
  den eigenen Browser verwenden 121

Anzeigen *(Forts.)*
Programmbeispiel 265
Verarbeitung auf linguistischer Basis 27
Arten von Textdokumenten 256
ASCII unstrukturiert, Dokumentart 256
ASCII, Dokumentart 256
Audio Extender 6
Auslöser
Beschreibung 12
erstellen 62

B

Basistextanalyse
Begriffe, die nichtalphanumerische Zeichen enthalten 21
für die Hervorhebung 27
Normalisierung 21
Satzerkennung 22
zum Indexieren von Begriffen 21
Bedarf an Plattenspeicherplatz für Indizes 17
Beenden einer Anzeigesitzung, API-Funktion
Beschreibung 207
Verwendung 123
Befehle
CHANGE INDEX SETTINGS 129
CHANGE TEXT CONFIGURATION 130
CONNECT 133
DB2TX 128
DELETE INDEX EVENTS 134
DISABLE DATABASE 135
DISABLE TEXT COLUMN 136
DISABLE TEXT FILES 137
DISABLE TEXT TABLE 138
ENABLE DATABASE 139
ENABLE TEXT COLUMN 139
ENABLE TEXT FILES 147
ENABLE TEXT TABLE 150
GET ENVIRONMENT 154
GET INDEX SETTINGS 154
GET INDEX STATUS 156
GET STATUS 157
GET TEXT CONFIGURATION 157
GET TEXT INFO 158
QUIT 158
REORGANIZE INDEX 159
RESET INDEX STATUS 160
TXICRT 163
TXIDROP 165
TXILIST 165
Befehle *(Forts.)*
TXNADD 166
TXNCHECK 167
TXNDROP 167
TXSAMPLE 168
TXSTART 168
TXSTATUS 169
TXSTOP 169
TXTHESC 169
TXTHESN 171
TXTRACE 172
TXVERIFY 178
UPDATE INDEX 160
Zusammenfassung, Client-Befehle 127
Zusammenfassung, Server-Befehle 163
Befehlszeilenprozessor
DB2TX, Befehl 128
Hilfe 128
QUIT, Befehl 158
starten 48
Begriffserweiterung für die Hervorhebung 27
Beispiel-API-Programm 265
Beispiel-UDFs
aktiv 95
Beispiel, Schnittstelle 43
Beispieltabellen
automatisch erstellen (TXSAMPLE) 55
Beschreibung 96
löschen 77
manuell erstellen 54
TXSAMPLE, Befehl 168
Bemerkungen 275
Benutzerdefinierte Funktionen
siehe UDFs
Boolesche Operatoren
& (AND) und | (OR) 103
NOT 108
Boolesches Suchargument 197
BOUND-Suche, Beispiel 109
BOUND, Schlüsselwort 198

C

CASE_ENABLED, Schlüsselwort
in ENABLE TEXT COLUMN 144
in ENABLE TEXT TABLE 151
CCSID
Beschreibung 259
Extrahieren aus einer Kennung 114
Funktion 182

CCSID *(Forts.)*
GET TEXT INFO, Befehl 158
Initialisieren in Kennungen 187
Liste 259
Schlüsselwort 133, 141
Standardwert in der Textkonfiguration 254
CHANGE INDEX SETTINGS, Befehl
Syntax 129
Verwendung 79
CHANGE TEXT CONFIGURATION, Befehl
Syntax 130
Verwendung 50
Client/Server-Umgebung 7
Codepage
siehe CCSID
COMMITCOUNT, Konfigurationsparameter
Beschreibung 63
in CHANGE TEXT CONFIGURATION 133
in ENABLE TEXT COLUMN 146
in ENABLE TEXT TABLE 161
Protokollbereich beibehalten 64
Standardwert in den Textkonfigurationseinstellungen 255
CONNECT, Befehl
Syntax 133
Verwendung 49
CONTAINS, Funktion
Beispiel 101
Syntax 182
COUNT, Schlüsselwort 196

D

Dateien außerhalb von DB2 UDB-Tabellen
siehe Externe Dateien
Dateikennung
siehe Kennung
Datenbank
aktivieren 56
CONNECT, Befehl 133
DISABLE DATABASE, Befehl 135
ENABLE DATABASE, Befehl 139
GET STATUS, Befehl 157
inaktivieren 76
sichern und zurückschreiben 93
Statusinformationen, anzeigen 82
Verbindung herstellen 49
Vorbereiten für die Suche 50
Datenstromsyntax 213
Datentypen von Textdokumenten
Funktion zum Umsetzen 66
unterstützt 141
DB2 Extender
Audio Extender 6
Beispiel für die Verwendung 4
Image Extender 5
Produktfamilie 4
Video Extender 5
DB2DBDFT, Umgebungsvariable 253
DB2INSTANCE, Umgebungsvariable 253
DB2TEXTFH, eindeutiger Typ 179
DB2TEXTH, eindeutiger Typ 179
DB2TEXTHLISTP, eindeutiger Typ 179
DB2TX_, Umgebungsvariablen
anzeigen 83
Beschreibung 253
DB2TX_INSTOWNER, Umgebungsvariable 253
DB2TX_INSTOWNERHOMEDIR, Umgebungsvariable 253
DB2TX, Befehlszeilenprozessor
Syntax 128
Verwendung 48
DB2TX.SAMPLE, Tabelle
Beschreibung 96
Dienstprogramm zum Erstellen 55
erstellen 54
löschen 77
Dekomposition von Komposita 23
DELETE INDEX EVENTS, Befehl
Beispiel 80
Syntax 134
DES_BROWSE, Option in DesGetSearchResultTable 217
DES_EXT.H, Kopfdatei 117
DES_EXTENDED, Option in DesOpenDocument 220
DES_FAST, Option in DesOpenDocument 220
DES_MATCH, Option in DesGetSearchResultTable 217
DES_NOBROWSE, Option in DesGetSearchResultTable 217
DES_RANK, Option in DesGetSearchResultTable 217
DES_RANKANDMATCH, Option in DesGetSearchResultTable 217
DES_TEXTHANDLEONLY, Option in DesGetSearchResultTable 217
DesCloseDocument, Funktion
Beschreibung 206
Verwendung 123

DesEndBrowseSession, Funktion
  Beschreibung 207
  Verwendung 123
DesFreeBrowseInfo, Funktion
  Beschreibung 208
  Verwendung 123
DesGetBrowseInfo, Funktion
  Beschreibung 208
  Verwendung 121
DesGetMatches, Funktion
  Beschreibung 211
  Verwendung 122
DesGetSearchResultTable, Funktion
  Beschreibung 216
  Verwendung 120
desmodel.ini 71
DesOpenDocument, Funktion
  Beschreibung 220
  Verwendung 122
DESRESTB, zum Erstellen einer Ergebnistabelle 218
DESSAMP1, Beispielprogramm 265
DesStartBrowseSession, Funktion
  Beschreibung 222
  Verwendung 121
DIRECTORY, Schlüsselwort
  Anzeigen der aktuellen Einstellung 87
  in CHANGE TEXT CONFIGURATION 132
  in ENABLE TEXT COLUMN 145
  in ENABLE TEXT TABLE 153
  Standardwert in den Textkonfigurationseinstellungen 255
DISABLE DATABASE, Befehl
  Syntax 135
  Verwendung 76
DISABLE TEXT COLUMN, Befehl
  Syntax 136
  Verwendung 74
DISABLE TEXT FILES, Befehl
  Syntax 137
  Verwendung 75
DISABLE TEXT TABLE, Befehl
  Syntax 138
  Verwendung 76
Dokument
  Anzeigen der Einstellungen 88
  Beschreibung 256
  CCSID 259
  Format in CHANGE TEXT CONFIGURATION 133
  Format in ENABLE TEXT COLUMN 142
  Format, Beschreibung 256
Dokument *(Forts.)*
  GET TEXT INFO, Befehl 158
  indexieren 11
  Informationen 88
  Sprache 258
  Struktur 68
  Umsetzen von Datentypen 66
  unterstützte Datentypen 141
  Vorbereiten für die Suche 50
Dokumentmodell
  Ändern der Dokumentmodelldatei 53
  Beschreibung 68
  MODEL, Schlüsselwort in der Suchsyntax 195
  SECTION, Schlüsselwort in der Suchsyntax 195
Dualer Index
  Beschreibung 15
  Suchoption, Standardwerte 198

E

Eindeutige Typen 179
Einschränkungen für Suchargumente 203
ENABLE DATABASE, Befehl
  Syntax 139
  Verwendung 56
ENABLE TEXT COLUMN, Befehl
  Syntax 139
  Verwendung 60
ENABLE TEXT FILES, Befehl
  Syntax 147
ENABLE TEXT TABLE, Befehl
  Syntax 150
  Verwendung 57, 72
Ereignis, Ursachencodes 267
Ergebnistabelle 218
Erkennen von Sätzen 22
Erstellen einer Beispieldatenbank 178
Erstellen von Beispieltabellen
  manuell 54
  TXSAMPLE, Befehl 168
Erweiterte Übereinstimmung 28
Erweiterung von Begriffen für die Hervorhebung 27
Escape-Zeichen
  Syntax 203
  Verwendung 105
Exemplar löschen
  TXIDROP, Befehl 165
  Vorgehensweise 44
Exemplare
  Befehl für Liste, TXILIST 165

Exemplare *(Forts.)*
erstellen 43
löschen 43
Umgebungsvariablen 253
EXPAND, Schlüsselwort 196
Extender
Audio Extender 6
Beispiel für die Verwendung 4
Image Extender 5
Produktfamilie 4
Video Extender 5
Externe Dateien
Abrufen oder Ändern eines Dateinamens in einer Kennung 183
aktivieren 72
Ändern des Pfades/Namens in Kennungen 114
Aspekte der Indexaktualisierung 78
DISABLE TEXT FILES, Befehl 137
ENABLE TEXT FILES, Befehl 147
Extrahieren des Pfades/Namens aus einer Kennung 114
FILE, Funktion 183
inaktivieren 75
Initialisieren von Kennungen 187
Kennungen 99

F

Fehlerereignisse
anzeigen 86
aufzeichnen 62
DELETE INDEX EVENTS 134
GET INDEX STATUS, Befehl 156
löschen 80
Ursachencodes 267
Fehlersuche 91
Feinanpassung einer vorherigen Suche 111
FFT, Dokumentart 256
FILE, Funktion
Beispiel 114
Syntax 183
Format von Textdokumenten
Ändern in Kennungen 114
Beschreibung 256
Extrahieren aus einer Kennung 114
FORMAT, Funktion 184
FORMAT, Schlüsselwort 133, 142
GET TEXT INFO, Befehl 158
in CHANGE TEXT CONFIGURATION 133
in ENABLE TEXT COLUMN 142
Format von Textdokumenten *(Forts.)*
Initialisieren in Kennungen 187
Standardwert in der Textkonfiguration 254
FORMAT, Funktion
Beispiel 114
Syntax 184
Freitextsuche
Beispiel 111
FUNCTION, Schlüsselwort
Beschreibung 66
in ENABLE TEXT COLUMN 141
Funktion
API-Funktionen 117
siehe auch Anwendungsprogrammierschnittstelle (API)
benutzerdefinierte Funktionen (UDFs) 95
Pfad für UDFs setzen 99
SET CURRENT FUNCTION PATH, Anweisung 99
zum Umsetzen von Datentypen 66
FUZZY FORM OF, Schlüsselwort 198

G

GET ENVIRONMENT, Befehl
Beispiel und Ausgabe 83
Syntax 154
GET INDEX SETTINGS, Befehl
Beispiel und Ausgabe 87
Syntax 154
GET INDEX STATUS, Befehl
Beispiel und Ausgabe 84
Syntax 156
GET STATUS, Befehl
Beispiel und Ausgabe 82
Syntax 157
GET TEXT CONFIGURATION, Befehl
Beispiel und Ausgabe 84
Syntax 157
GET TEXT INFO, Befehl
Beispiel und Ausgabe 88
Syntax 158
Große Tabellen, aktivieren 63
Grundform, Begriffe reduzieren 22
GUI, Beispiel 43

H

HANDLE_LIST, Funktion
Syntax 186
Verwendung 115

HANDLE, Funktion
  Syntax 185
  Verwendung 115
Hervorhebungsinformationen
  Datenstrom 122
  Datenstromsyntax 211
Hilfe für Befehle 128
HTML, Dokumentart 256
Hybridsuche, Beispiel 111

I

ID des codierten Zeichensatzes
  siehe CCSID
Image Extender 5
IN SAME PARAGRAPH AS, Schlüsselwort 197
IN SAME SENTENCE AS, Schlüsselwort 197
Include-Datei, des_ext.h 117
Index
  Aktualisieren für externe Dateien 78
  Aktualisierungsfrequenz 262
  Ändern der Aktualisierungsfrequenz 79
  Ändern der aktuellen Einstellungen 79
  Ändern des Indextyps 16
  Anzeigen der aktuellen Einstellungen 87
  CASE_ENABLED, Option 15
  CHANGE INDEX SETTINGS, Befehl 129
  CHANGE TEXT CONFIGURATION, Befehl 130
  dual 15
  Erstellen von verschiedenen Typen für eine Textspalte 63
  GET INDEX SETTINGS, Befehl 154
  GET INDEX STATUS, Befehl 156
  GET TEXT CONFIGURATION, Befehl 157
  Größenberechnung 17
  INDEXOPTION in ENABLE TEXT COLUMN 143
  INDEXOPTION in ENABLE TEXT TABLE 151
  INDEXTYPE in CHANGE TEXT CONFIGURATION 131
  INDEXTYPE in ENABLE TEXT COLUMN 143
  INDEXTYPE in ENABLE TEXT TABLE 151
  linguistisch 13
  Ngram 15
  planen 11
  präzise 14
  regelmäßige Indexaktualisierung 262
  reorganisieren 81
  sichern und zurückschreiben 93
  sofortige Indexaktualisierung 77
  Standardtyp in den Textkonfigurationseinstellungen 255
Index *(Forts.)*
  Tabelle mit einem gemeinsam benutzten Index 16
  TABLESPACE in CHANGE TEXT CONFIGURATION 132
  Textkonfiguration ändern 50
  Typen 13
  Überblick 11
  UPDATE INDEX, Befehl 160
  verwalten 77
Indexieren, Verarbeitung auf linguistischer Basis 19
Indexierungsereignisse
  Ursachencodes 267
Indexierungsereignisse, löschen
  Beispiel 80
  Syntax 134
Indexmerkmale
  anzeigen 87
  in ENABLE TEXT COLUMN 140
  in ENABLE TEXT FILES 148
  in ENABLE TEXT TABLE 150
  Standardwerte in den Textkonfigurationseinstellungen 255
INDEXOPTION, Schlüsselwort
  in CHANGE TEXT CONFIGURATION 132
  in ENABLE TEXT COLUMN 143
  in ENABLE TEXT TABLE 151
INDEXPROPERTY, Schlüsselwort
  in ENABLE TEXT COLUMN 144
  in ENABLE TEXT TABLE 152
Indexstatus, anzeigen
  Beispiel und Ausgabe 84
  Syntax 156
Indexstatus, zurücksetzen
  Beispiel 79
  Syntax 160
Indextyp, ändern
  ändern 16
  Erstellen von verschiedenen Typen für eine Textspalte 63
INDEXTYPE, Schlüsselwort
  in CHANGE TEXT CONFIGURATION 131
  in ENABLE TEXT COLUMN 143
  in ENABLE TEXT TABLE 151
Informationen zu Textdokumenten
  Anzeigen der aktuellen Einstellung 88
  Arten 256
  CCSID 259
  Format 256
  GET TEXT INFO, Befehl 158
  Sprache 258

INIT_TEXT_HANDLE, Funktion
  Beispiel 113
  Syntax 187
Initialisieren einer Kennung
  INIT_TEXT_HANDLE, Funktion 187
  Vorgehensweise 113

K

Katalogsicht
  erstellen 56
  Inhalt 89
  löschen 77
Kennung
  Abrufen aus einer Liste von Kennungen 185
  Ändern von Format und Sprache 114
  Beschreibung 97
  CCSID, Funktion 182
  eindeutiger Typ, DB2TEXTFH 179
  eindeutiger Typ, DB2TEXTH 179
  Extrahieren von CCSID, Format und Sprache 114
  FORMAT, Funktion 184
  für externe Dateien 99
  initialisieren 113
  LANGUAGE, Funktion 188
  Setzen und Extrahieren von Informationen 113
  Verwendung von Listen zur Verbesserung der Leistung 115
Kennungslistenzeiger (eindeutiger Typ DB2TEXTHLISTP) 179
Knoten
  Befehl zum Hinzufügen von Knoten, TXNADD 166
  Befehl zum Löschen von Knoten, TXNDROP 167
  Befehl zum Prüfen von Knoten, TXNCHECK 167
  Umgebung mit mehreren Knoten 67
Knoten, zuordnen 9
Knotengruppen und Tabellenbereiche 62
Kompilieren einer Ngram-Thesaurus-Definitionsdatei 171
Kompilieren einer Thesaurus-Definitionsdatei 169
Komposita, trennen 23
Konfiguration 253
Konfigurationstabelle
  anzeigen 84
  CHANGE TEXT CONFIGURATION, Befehl 130
  erstellen 56
  GET TEXT CONFIGURATION, Befehl 157
Kopfdatei, des_ext.h 117

L

LANGUAGE, Funktion
  Beispiel 114
  Syntax 188
LANGUAGE, Schlüsselwort 133, 141
Lauterweiterung
  Beispiel 109
  Beschreibung 26
Leistung, verbessern 115
Linguistischer Index
  Beschreibung 13
  Suchoption, Standardwerte 198
Logischer Knoten, zuordnen 9
LOGPRIMARY, LOGSECOND und LOGFILSIZ, Parameter in DB2 UDB 63

M

Mehrpartitionsknotengruppe 9
Microsoft, Dokumentart 256

N

Nach Text suchen
  Abrufen der Anzahl der gefundenen Übereinstimmungen 102
  Abrufen der Rangordnung eines gefundenen Dokuments 102
  Durchführen einer Abfrage 101
  Erstellen einer Liste von gefundenen Dokumenten 186
  Feinanpassung einer vorherigen Suche 111
  Programmbeispiel 265
  REFINE, Funktion 190
  SEARCH_RESULT, Funktion 190
  Syntax 194
  Überblick 100
  Verbessern der Leistung 115
  Verwendung der API 120
Nachrichten 235
Ngram-Index
  Beschreibung 15
  CASE_ENABLED, Option 15
  Suchoption, Standardwerte 198
NO_OF_DOCUMENTS, Funktion
  Syntax 189
NODE, Schlüsselwort
  in ENABLE TEXT COLUMN 145
  in ENABLE TEXT TABLE 153

Normalisierung von Begriffen 21
NORMALIZED, Schlüsselwort
in ENABLE TEXT COLUMN 144
NOT
Boolescher Operator 108
Schlüsselwort in einem Suchargument 197
NUMBER_OF_MATCHES, Funktion, Syntax 189

O

Öffnen eines Dokuments, API-Funktion
Beschreibung 220
Verwendung 122
ON NODE, Schlüsselwort
in ENABLE TEXT COLUMN 145
in ENABLE TEXT TABLE 153
OR, Boolescher Operator 103

P

Parallelismus, Konzepte 9
Physischer Knoten, zuordnen 9
Planen eines Textindexes 11
Plattenspeicherplatz für Indizes 17
Platzhalterzeichen
in einem Suchbegriff 104
Platzhalterzeichen für Wörter 26
Verarbeitung auf linguistischer Basis 26
Platzhalterzeichen für Zeichen 26
Präziser Index
Beschreibung 14
Suchoption, Standardwerte 198
PRECISE FORM OF, Schlüsselwort 198
Protokollbereich, nicht genügend 63
Protokolltabelle
Beschreibung 12
erstellen 62
Extrahieren von Fehlerereignissen 86
Zuordnen zum Tabellenbereich 62

Q

QUIT, Befehl
Syntax 158
Verwendung 74

R

Rangordnung
aus DesGetSearchResultTable 120
in einem Suchergebnis 102
RANK, Funktion
Beispiel 102
Syntax 190
Reduzieren von Begriffen auf ihre Grundform 22
REFINE, Funktion
Beispiel 111
Syntax 190
Regelmäßige Indexaktualisierung
siehe Aktualisieren eines Textindexes
Regeln für Suchargumente 203
REORGANIZE INDEX, Befehl
Beispiel 81
Syntax 159
RESET INDEX STATUS, Befehl
Beispiel 79
Syntax 160
RESULT LIMIT, Schlüsselwort 196
Rückkehrcodes, von der API 225

S

Satzerkennung 22
Satztrennung 14
Schließen eines Dokuments, API-Funktion
Beschreibung 206
Verwendung 123
Schnittstelle, Beispiel 43
SEARCH_RESULT, Funktion
Beispiel 115
Syntax 190
Server
Anzeigen des Status 169
starten 168
stoppen 169
TXICRT, Befehl 163
TXIDROP, Befehl 165
TXILIST, Befehl 165
TXNADD, Befehl 166
TXNCHECK, Befehl 167
TXNDROP, Befehl 167
TXSAMPLE, Befehl 168
TXSTART, Befehl 168
TXSTATUS, Befehl 169
TXSTOP, Befehl 169
TXTRACE, Befehl 172

Server *(Forts.)*
TXVERIFY, Befehl 178
SET CURRENT FUNCTION PATH, Anweisung 99
Shell-Profile 253
Sichern und Zurückschreiben 93
Spalte
aktivieren 60
Aktivieren für verschiedene Indextypen 63
Aktivieren in einer großen Tabelle 63
DISABLE TEXT COLUMN, Befehl 136
ENABLE TEXT COLUMN, Befehl 139
inaktivieren 74
Speicherplatz freigeben für Anzeigeinformationen, API-Funktion
Beschreibung 208
Verwendung 123
Sprache von Textdokumenten
Ändern in Kennungen 114
Beschreibung 258
Extrahieren aus einer Kennung 114
GET TEXT INFO, Befehl 158
in einem Suchbegriff 107
Initialisieren in Kennungen 187
LANGUAGE, Funktion 188
LANGUAGE, Schlüsselwort 133, 141
linguistische Funktionen 30
Liste 258
Standardwert in der Textkonfiguration 254
SQL-Datentypen
siehe Datentypen von Textdokumenten
SQLSTATES, die von UDFs zurückgegeben werden 235
Starten des Text Extender-Servers 168
Starten einer Anzeigesitzung, API-Funktion
Beschreibung 222
Verwendung 121
Status des Text Extender-Servers 169
Status eines Indexes
anzeigen 84
Anzeigen des aktuellen Status 156
zurücksetzen 79
zurücksetzen nach einem Fehler 79
STEMMED FORM OF, Schlüsselwort 198
Stoppen des Text Extender-Servers 169
Stoppwörter
als Teil der Basistextanalyse 23
Beschreibung 11
Editieren einer Stoppwortdatei 52
Listen 31
Struktur von Dokumenten
Aktivieren der Abschnittsunterstützung 68
MODEL, Schlüsselwort in der Suchsyntax 195
SECTION, Schlüsselwort in der Suchsyntax 195
Suchbeispiel 106
Suchargument
| (OR), Operator 197
& (AND), Operator 197
AND, Schlüsselwort 197
angeben 102
Beschreibung 193
BOUND-Suche 109
BOUND, Schlüsselwort 198
COUNT, Schlüsselwort 196
EXPAND, Schlüsselwort 196
Freitextsuche 111
FUZZY FORM OF, Schlüsselwort 198
Hybridsuche 111
IN SAME PARAGRAPH AS 197
IN SAME SENTENCE AS 197
MODEL, Schlüsselwort 195
NOT, Schlüsselwort 197
PRECISE FORM OF, Schlüsselwort 198
RESULT LIMIT, Schlüsselwort 196
SECTION, Schlüsselwort 195
STEMMED FORM OF, Schlüsselwort 198
Suche nach grober Übereinstimmung 109, 198
Suchen mit & und | 103
Suchen mit NOT 108
Suchen nach ähnlich klingenden Wörtern 109
Suchen nach Begriffen im selben Absatz 106
Suchen nach Begriffen im selben Satz 106
Suchen nach Begriffen in beliebiger Reihenfolge 106
Suchen nach Begriffen in Dokumentabschnitten 106
Suchen nach Begriffen in verschiedenen Sprachen 107
Suchen nach mehreren Begriffen 102
Suchen nach Synonymen 106
Suchen nach Teilen eines Begriffs 104
Suchen nach Variationen eines Begriffs 104
SYNONYM FORM OF, Schlüsselwort 198
Syntax 194
TERM OF, Schlüsselwort 196
Thesaurus-Suche 110
THESAURUS, Schlüsselwort 196
unter Verwendung von Platzhalterzeichen 104
Zusammenfassung der Regeln und Einschränkungen 203

Suche nach grober Übereinstimmung, Beispiel 109
Suchfunktionen 117
Suchstatus, anzeigen
  Beispiel und Ausgabe 84
  Syntax 156
Suchstatus, zurücksetzen
  Beispiel 79
  Syntax 160
Synchronisieren eines Indexes mit der Datenbank
  siehe Aktualisieren eines Textindexes
Synonyme
  Beschreibung 25
  in einem Suchargument 106
  SYNONYM FORM OF, Schlüsselwort 198

T

Tabelle
  siehe Texttabelle
Tabelle mit einem gemeinsam benutzten Index
  Beschreibung 16, 17
  ENABLE TEXT TABLE, Befehl 150
  erstellen 57
Tabellenbereich 62
Tabellenbereiche und Knotengruppen 62
TERM OF, Schlüsselwort 196
Text Extender-Exemplar erstellen
  TXICRT, Befehl 163
Textdokument
  siehe Dokument
Textindex
  siehe Index
TEXTINDEXES, Katalogsicht
  erstellen 56
  Inhalt 89
  löschen 77
Textkennung
  siehe Kennung
Textkonfigurationseinstellungen
  ändern 50
  anzeigen 84
  Installationstandardwerte 254
Textmerkmale
  Beschreibung 256
  CCSID 259
  Format 256
  in ENABLE TEXT COLUMN 140
  in ENABLE TEXT FILES 148
  Sprache 258
  Standardwerte in der Textkonfiguration 254
Textspalte
  siehe Spalte
Texttabelle
  Aktivieren einer Spalte in einer großen Tabelle 63
  DISABLE TEXT TABLE, Befehl 138
  ENABLE TEXT TABLE, Befehl 150
  sichern und zurückschreiben 93
Thesaurus-Suche
  Beispiel 110
  Erstellen eines Ngram-Thesaurus 41
  Erstellen eines Thesaurus 37
  Kompilieren einer
    Ngram-Thesaurus-Definitionsdatei 171
  Kompilieren einer Thesaurus-Definitionsdatei 169
  Konzepte 32
  Syntax 196
  THESAURUS, Schlüsselwort 196
  TXTHESC, Befehl 169
  TXTHESN, Befehl 171
Tiefe von Begriffen in einem Thesaurus, angeben 196
Trace-Funktion
  einstellen 91
  TXTRACE, Befehl 172
TXICRT, Befehl
  Syntax 163
  Text Extender-Exemplar erstellen 43
TXIDROP, Befehl
  Syntax 165
TXILIST, Befehl
  Syntax 165
TXNADD, Befehl
  Syntax 166
TXNCHECK, Befehl
  Syntax 167
TXNDROP, Befehl
  Syntax 167
TXSAMPLE, Befehl
  Syntax 168
  Verwendung 55
TXSAMPLE.UDF
  aktiv 95
TXSTART, Befehl
  Syntax 168
  Verwendung 44
TXSTATUS, Befehl
  Syntax 169
  Verwendung 44
TXSTOP, Befehl
  Syntax 169
  Verwendung 44

TXTHESC, Befehl
Syntax 169
TXTHESN, Befehl
Syntax 171
TXTRACE, Befehl
Syntax 172
Verwendung 91
TXVERIFY
Erstellen einer Beispieldatenbank 55
TXVERIFY, Befehl
Syntax 178
Typen von Textindizes
CASE_ENABLED, Option 15
CHANGE TEXT CONFIGURATION, Befehl 131
dual 15
GET INDEX SETTINGS, Befehl 154
INDEXTYPE in CHANGE TEXT CONFIGURATION 131
INDEXTYPE in ENABLE TEXT COLUMN 143
INDEXTYPE in ENABLE TEXT TABLE 151
linguistisch 13
Ngram 15
präzise 14
Standardwert in den Textkonfigurationseinstellungen 255
Suchoption, Standardwerte 198

U

Überblick über die DB2 Extender 3
Übereinstimmung
aus DesGetSearchResultTable 120
DesGetMatches, Funktion 211
in einem Suchergebnis 102
NUMBER_OF_MATCHES, Funktion 189
Übereinstimmung, erweitert 28
UDFs
Angeben von Suchargumenten 102
Beschreibung 95
CCSID 182
CONTAINS 182
Feinanpassung einer vorherigen Suche 111
FILE 183
FORMAT 184
Funktionspfad 99
HANDLE 185
HANDLE_LIST 186
INIT_TEXT_HANDLE 187
LANGUAGE 188
NO_OF_DOCUMENTS 189
UDFs *(Forts.)*
NUMBER_OF_MATCHES 189
RANK 190
Referenz 179
REFINE 190
SEARCH_RESULT 190
Setzen und Extrahieren von Informationen in Kennungen 113
Suchen nach Text 100
Überblick 181
Verbessern der Suchleistung 115
zurückgegebene SQLSTATES 235
UDTs 179
Umgebung mit mehreren Knoten 67
Umgebung, Client/Server 7
Umgebungsvariablen 253
anzeigen 83
Beschreibung 253
GET ENVIRONMENT, Befehl 154
Unterstützung partitionierter Datenbanken 9
UPDATE INDEX, Befehl
Beispiel 77
Syntax 160
UPDATEFREQ, Schlüsselwort
in CHANGE INDEX SETTINGS 130
in CHANGE TEXT CONFIGURATION 132
in ENABLE TEXT COLUMN 145
in ENABLE TEXT TABLE 152
UPDATEINDEX, Schlüsselwort
Anzeigen der aktuellen Einstellung 87
GET INDEX SETTINGS, Befehl 154
in CHANGE TEXT CONFIGURATION 132
in ENABLE TEXT COLUMN 145
Standardwert in den Textkonfigurationseinstellungen 255

V

Variablen
Anzeigen von Umgebungsvariablen 83
Beschreibung der Umgebungsvariablen 253
GET ENVIRONMENT, Befehl 154
Verarbeitung auf linguistischer Basis
Basistextanalyse 21
Begriffserweiterung 27
beim Indexieren 19
Beschreibung 19
erweiterte Übereinstimmung 28
für die Abfrage 24
für die Anzeige 27

Verarbeitung auf linguistischer Basis *(Forts.)*
für die unterstützten Sprachen 30
Lauterweiterung 26
Platzhalterzeichen 26
Platzhalterzeichen für Wörter 26
Platzhalterzeichen für Zeichen 26
Reduzieren von Begriffen auf ihre Grundform 22
Stoppwortfiltern 23
Synonyme 25
Trennen von Komposita 23
Verarbeitungsmerkmale
Standardwerte in den Textkonfigurationseinstellungen 255
Verbindung zu einer Datenbank herstellen
CONNECT, Befehl 133
Vorgehensweise 49
Verwaltung
Abkürzungsdatei, editieren 52
Aktivieren einer Datenbank 56
Aktivieren einer Textspalte 60
Aktivieren einer Texttabelle 57
Aktivieren von externen Dateien 72
Aktualisieren eines Indexes für externe Dateien 78
Ändern der Indexeinstellungen 79
Ändern von Stoppwort- und Abkürzungsdateien 52
Anzeigen der Einstellungen für Textdaten 88
Anzeigen der Indexeinstellungen 87
Anzeigen des Indexstatus 84
Anzeigen des Status von Datenbank, Tabelle und Spalte 82
Beenden einer Verwaltungssitzung 74
Befehlszeilenprozessor 128
Befehlszeilenprozessor starten 48
Befehlszusammenfassung, Client 127
Befehlszusammenfassung, Server 163
Beispiel, Wizard 43
Beispieldatenbank, Dienstprogramm zum Erstellen 55
Beispieltabellen, erstellen 54
CHANGE INDEX SETTINGS, Befehl 129
CHANGE TEXT CONFIGURATION, Befehl 130
CONNECT, Befehl 133
Datenbank für die Suche vorbereiten 50
DB2TX, Befehl 128
DELETE INDEX EVENTS 134
DELETE INDEX EVENTS, Befehl 134
DISABLE DATABASE, Befehl 135
DISABLE TEXT COLUMN, Befehl 136
DISABLE TEXT FILES, Befehl 137
DISABLE TEXT TABLE, Befehl 138
Verwaltung *(Forts.)*
ENABLE DATABASE, Befehl 139
ENABLE TEXT COLUMN, Befehl 139
ENABLE TEXT FILES, Befehl 147
ENABLE TEXT TABLE, Befehl 150
Erstellen einer Beispieldatenbank 178
Erstellen von Beispieltabellen 54
GET ENVIRONMENT, Befehl 154
GET INDEX SETTINGS, Befehl 154
GET INDEX STATUS, Befehl 156
GET STATUS, Befehl 157
GET TEXT CONFIGURATION, Befehl 157
GET TEXT INFO, Befehl 158
GUI, Beispiel 43
Inaktivieren einer Datenbank 76
Inaktivieren einer Textspalte 74
Inaktivieren einer Texttabelle 76
Inaktivieren von Textdateien 75
Kompilieren einer Ngram-Thesaurus-Definitionsdatei 171
Kompilieren einer Thesaurus-Definitionsdatei 169
Löschen von Indexereignissen 80
QUIT, Befehl 158
Reorganisieren eines Indexes 81
REORGANIZE INDEX, Befehl 159
RESET INDEX STATUS, Befehl 160
Server-Status anzeigen 44
sichern und zurückschreiben 93
sofortiges Aktualisieren eines Indexes 77
Starten des Text Extender-Servers 168
Statusinformationen, abrufen 82
Stoppen des Text Extender-Servers 169
Stoppwortdatei, ändern 52
Text Extender-Exemplar erstellen 43
Text Extender-Exemplar löschen 43
Textkonfiguration ändern 50
Trace-Funktion 91
TXICRT, Befehl 163
TXIDROP, Befehl 165
TXILIST, Befehl 165
TXNADD, Befehl 166
TXNCHECK, Befehl 167
TXNDROP, Befehl 167
TXSAMPLE, Befehl 168
TXSTART, Befehl 168
TXSTATUS, Befehl 169
TXSTOP, Befehl 169
TXTHESC, Befehl 169
TXTHESN, Befehl 171
TXTRACE, Befehl 172

Verwaltung *(Forts.)*
TXVERIFY, Befehl 178
Umgebungsvariablen 253
UPDATE INDEX, Befehl 160
Verbindung zur Datenbank herstellen 49
Verwalten von Textindizes 77
Verwaltungssitzung starten 48
Verwaltungsüberblick 45
Wizard, Beispiel 43
Zurücksetzen des Indexstatus 79
Zusammenfassung der Befehle, Client 127
Zusammenfassung der Befehle, Server 163
Verzeichnis für den Index
GET INDEX SETTINGS, Befehl 154
Video Extender 5
Vorkommen eines Suchbegriffs 189

W

Wizard, Beispiel 43
WordPerfect, Dokumentart 256
Wörterverzeichnisdateinamen 31
Worttrennung 14

Z

Zugriffsfunktion
Beschreibung 66
in ENABLE TEXT COLUMN 141



Printed in U.S.A.

SC12-2709-00

